Masaryk University, Brno, Czech Republic

Proceedings of the European conference TOWARDS *e*ENVIRONMENT

Opportunities of SEIS and SISE:
Integrating Environmental Knowledge in Europe

Chief Editor:
Jiří Hřebíček

Main Editors:
Jiří Hradec, Emil Pelikán, Ondřej Mirovský,
Werner Pillmann, Ivan Holoubek, Thomas Bandholtz

European conference of the Czech Presidency of the Council of the EU

TOWARDS *e*ENVIRONMENT

March 25–27, 2009 | Prague, Czech Republic

Partners

Sponsors

General partners

Main partners

Partners

Media partners

A member of the Editorial Board and one anonymous referee refereed each session paper in these proceedings. A member of the Editorial Board refereed each workshop paper/abstract.

ISBN 978-80-210-4824-9

Committees

Chairman

Prof. Dr. Jiří Hřebíček,
Masaryk University, Czech Republic

Co-organiser

Zuzana Brychová
GUARANT International spol. s r.o.

Scientific Committee

Thomas Bandholtz innoQ Deutschland GmbH, DE
Vladimir Benko Slovak Environmental Agency, SK
Vladislav Bízek Technology Centre AS CR, CZ
Thomas Buchsbaum CAHDE, Council of Europe, AT
Jose Mario Caridad University of Cordoba, ES
Ralf Denzer Saarland University of Applied Sciences, DE
Werner Geiger Research Center Karlsruhe, DE
Albrecht Gnauck Brandenburg University of Technology Cottbus, DE
Jorge Marx Gómez Carl von Ossietzky University Oldenburg, DE
Giorgio Guariso Politecnico di Milano, IT
Lorenz Hilty EMPA, St. Gallen, CH
Ivan Holoubek Masaryk University, CZ
Jiří Hradec Czech Environmental Information Agency, CZ
Olgierd Hryniewicz Systems Research Institute Polish Academy of Sciences, PL
Jiří Hřebíček Masaryk University, CZ, ***chairman***
Karel Charvát Czech Centre for Science and Society, CZ
Ralf Isenmann Fraunhofer Institute for Systems and Innovation Research (ISI) Karlsruhe, DE
Stefan Jensen European Environment Agency, DK
Kostas Karatzas Aristotle University, GR
Arno Kaschl European Commission, DG Enterprise and industry, GMES Bureau, BE
Gerlinde Knetsch Federal Environmental Agency, DE
Jan Kolář Czech Space Office, CZ
Milan Konečný Masaryk University, CZ
Horst Kremers CODATA - Germany, DE
Daniel Vidal-Madjar Centre National d'Etudes Spatiales, FR
Colette Maloney European Commission, DG Enterprise and industry, GMES Bureau, BE
Ondřej Mirovský Czech Space Office, CZ
Jaroslav Myšiak FEEM, Venice, IT
Márta Nagy-Rothengass European Commission, DG Information Society and Media, BE
Patricia Nuttall Centre for Ecology & Hydrology, UK
Bernd Page University of Hamburg, DE
Meropi Paneli European Commission, DG Environment, BE
Emil Pelikán Institute of Computer Science of Academy of Science, CZ
Thomas Pick Ministry of Environment of Lower Saxony, DE
Werner Pillmann International Society for Environmental Protection, AT
Jaroslav Pokorný Charles University, CZ
Wolf-Fritz Riekert Stuttgart Media University, DE
Iva Ritschelová University of Jan Evagenlista Purkyně, CZ
Andrea Emilio Rizzoli IDSIA, CH

Bernt Röndell European Environment Agency, DK
Ivica Ruzic Rudjer Boskovic Institute, HR
Miquel Sànchez-Marrè Technical University of Catalonia, ES
Arno Scharl MODUL University Vienna, AT
Gerald Schimak ARC Seibersdorf research, AT
Martin Schreiber Leuphana Universität Lüneburg, DE
Michel Schouppe European Commission, DG Information Society and Media, BE
Jean Philippe Steyer LBE, INRA, France
Hartmut Streuff Federal Environment Ministry (BMU), DE
Julius Stuller Institute of Computer Science of Academy of Science, CZ
Alberto Susini Geneva Labour inspectorate (OCIRT), Switzerland
David Swayne University of Guelph, CA
Olga Štěpánková Czech Technical University, CZ
Nguyen Xuan Thinh University of Rostock, DE
Klaus Tochtermann Know-Center, Graz University of Technology, AT
Thomas Usländer Fraunhofer IITB, DE
Wilhelm Vogel Umweltbundesamt, AT
Kristina Voigt Helmholtz Zentrum Muenchen, German Research Center for Environmental Health (GmbH), DE
Irina Zálišová EPMA, CZ

Local Organising Committee

Jarmila Cikánková, Czech Environmental Information Agency
Eva Hillerová, Technology Centrum AV CR
Jiří Hradec, Czech Environmental Information Agency
Ondřej Mirovský, Czech Space Office
Emil Pelikán, Institute of Computer Science of Academy of Science
Lenka Vrtíšková, Ministry of Environment

Conference Administration

Zuzana Brychová
GUARANT International spol. s r.o.

Chairman Assistant

Lucie Pekárková
Masaryk University

Acknowledgements

The chief editor would like to acknowledge the work of main editors and all the members of the editorial board. Due to their wonderful work, it has been possible to provide a fair and thorough review to all the papers and abstracts in the proceedings.

The chairman would like to thank the great efforts of all the **Local Organising Committee members**. Without them, the European conference TOWARDS eENVIRONMENT would not have been possible. Thank you all for your collaboration.

Preface

Information and communication technology (ICT) plays an important role in the European and global economy, but ICT also embodies other possibilities to improve the well being of the citizen of the world. The member states of European Union (EU) suffer hard from the financial crisis and the raising unemployment in the Czech Presidency of the Council of EU, but do tremendous work to contribute to a sustainable future. Current environmental challenges are the world wide exorbitant resource consumption, air pollution, the loss of biodiversity, waste and most important for the world development: the climate change. Such main problem areas and especially environmental relevant degradations indicate the need to raise the level of a combined discussion on the economic, environmental and social objectives, in the sense of a sustainable development. A positive economic and social development including the preservation of the environment is a precondition to the development of our societies and eEnvironment will play the important role.

The idea of the eEnvironment (Electronic access to Environmental information) was presented at the second CAHDE (Ad hoc Committee on E-Democracy of the Council of Europe) plenary meeting, which was held in Strasbourg on 8 and 9 October 2007. The eEnvironment legal basis is the Aarhus Convention, which is implemented in the European Community and supported by the EU Directives: *2003/4/EC Public Access to Environmental Information; 2003/35/EC Public Participation; 2003/98/EC Re-use of Public Sector Information and 2007/2/EC Infrastructure for Spatial Information in the European Community (INSPIRE).* The eEnvironment belongs into eParticipation and eGovernment initiative and it is going to be one of the fundaments of eDemocracy. The embedding of FP7 project activities in the conference, e.g. from ICT-ENSURE, provides a good additional opportunity for networking between experts from different fields.

The Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions *COM(2008) 46 final: Towards a Shared Environmental Information System (SEIS)* provides the political, organization and technical basis for the eEnvironment, thus that any citizen to be informed about environmental matters and to use this information for active participation in decision making. The group *Go4*: the Directorate-General (DG) Environment; the European Environment Agency (EEA); the Statistical Office of the European Communities (Eurostat) and the Joint Research Centre (JRC), controls the development of the SEIS. It collaborates with *Global Monitoring for Environment and Security (GMES)* initiative, which represents a concerted effort to bring data and information providers together with users, so they can better understand each other and make environmental and security-related information available to the people who need it through new enhanced services.

The objective "ICT-2007.6.3: ICT for Environmental Management and Energy Efficiency" of EU's Seventh Framework Programme (FP7), introduced a *Single Information Space in Europe for the Environment (SISE)* in which environmental institutions, service providers and citizens can collaborate or use available information without technical restraints. This will be a platform also for an efficient research support of eEnvironment.

The European conference of the Czech Presidency of the EU Council TOWARDS eENVIRONMENT streamlines a development of the eEnvironment networks - the future Internet, the achievement of interoperability, the extension of digital contents, the support of new ways of mobility and the development of a cutting-edge technical basis for integrating information are tasks for research and enterprises.

There is one duty of the TOWARDS eENVIRONMENT conference formulated in its subtitle *Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe* to build a SEIS, a system, which integrates information on air quality, freshwater, climate effects, nature conservation, waste management and all the important issues, where especially DG Environment and EEA deal with in an multidisciplinary manner. The framework of the SISE could support to build SEIS. The SISE is an ICT research vision for real-time connectivity between multiple environmental resources which would allow seamless cross-system search, as well as cross-border, multi-scale, multi-disciplinary data acquisition, pooling and sharing. Furthermore, it would allow for service-chaining on the Web, thereby stimulating data integration into innovative value-added Web services, which needs SEIS.

Merge interests of DG Environment, DG Enterprise and Transport, and DG Information Society and Media the TOWARDS eENVIRONMENT conference is dedicated to information exchange among public administrations, EU institutions, environmental agencies, scientists and businesses involved in development and use of environmental informatics for the delivery of modern eEnvironment services in Europe.

The conference sessions are focused on the following topics:

- **Session S1**: ICT research towards a Single Information Space in Europe for the Environment. How can ICT contribute to strengthen a collaborative information space on the web that will maximise the use of distributed environmental information for multiple purposes and actors?
- **Session S2**: Building SEIS, Europe's backbone for eGovernment environmental services. Does Europe champion challenges and seize the opportunities in provision of eEnvironment/eParticipation services?
- **Session S3**: ICT for monitoring and control of energy efficiency and security, modelling of air pollution, climate changes; for support of decision making in environment protection,
- **Session S4**: The Global Monitoring for Environment and Security (GMES) and its complex EU dimension: GMES and European space agency; GMES as an efficient European tool harmonise data collection, provision and final use; GMES services for end users from New EU member countries; atmosphere and land services for users.

The conference workshops are devoted to the following topics:

- **Workshop ICT-ENSURE** – European ICT Environmental Sustainability Research. This workshop is the third in a series of regional workshops across Europe. Its special focus in Prague is on ICT supporting environmental sustainability in Europe. The workshop targets researchers whose fields of interest address specifically the enabling role of ICT in SEIS and SISE and in the associated sustainability domains such as Energy consumption/efficiency, climate change, eco-industrial applications, industrial ecology, agriculture, landscape and biodiversity, personalised information services and quality of life, sustainable urban development, health and environmental risk management.
- **Workshop eEcoRA** – Information Technologies in Ecological Risk Assessment, This workshop summarizes current methodology of ecological risk assessment (EcoRA) in view of most common problems with data accessibility and processing. Main part of the workshop is focused on processing and modelling of complex abiotic and biotic data entering EcoRA and their aggregation according to rules given by problem to be solved. All issues are supported by practical case studies working with real data from bio-monitoring of fresh waters and soil ecosystems.
- **Workshop eEnvironment Terminology**. This workshop presents several domain-specific and interdisciplinary examples and discusses common design issues such as terminology structure models, cross-referencing, symmetric vs. asymmetric multilingualism, identity and reference, publishing terminology in the Web, and linking environmental data to such published terminology.
- **Workshop SEIS/NESIS** - ICT Components for SEIS around Europe. The NESIS is a Network promoted within the ICT-PSP program with the aim to promote the uptake of ICT solutions by public authorities in providing information for the monitoring and reporting of environmental impacts and threats. This workshop presents the results of the assessment of the status of environmental monitoring and reporting in the countries of the network partners based on a draft version of a Summary Report and good practices in ICT solution in handling environmental data, as already collected by the project partners into the NESIS Best Practice Catalogue.

Several contributions of sessions (e.g. session S4) and workshops (e.g. workshop W5) were presented only orally at the conference and their power point presentations you can download from the conference web http://www.e-envi2009.org/presentation.

The editors hope you will find new knowledge and ideas in the conference proceedings and web, where you can meet reviewed contributions of authors and experts on leading edge ICT technologies for environment of the above mentioned topics.

The editors are grateful to all authors for their valuable contributions to the conference and the members of Programme Committee for reviewing more than 150 submitted abstracts. Special acknowledgements belong to representatives of *DG Environment, DG Enterprise and Transport*, and *DG Information Society and Media* of the Commission, the *European Environment Agency*, and the *European Space Agency* for their kind support and useful advices with respect to the improvement of organisation of TOWARDS eENVIRONMENT conference.

We would like also to thank the Masaryk University, its Institute of Biostatistics and Analyses, the Ministry of Environment of the Czech Republic, the Czech Environmental Information Agency, the Czech Space Office, the Institute of Computer Science of the Academy of Science of the Czech Republic (AS CR) and Technology Centre AS CR for the scientific and organisational support of conference, the company Guarant International and Mrs.

Zuzana Brychová for the administrative support of conference and Miroslav Kubásek and Lucie Pekárková - the conference web administration and editorial team for their enormous work in the editing web and proceedings.

Last but not the least we thank all partners of conference under their sponsorship and auspices is the conference hold for their valuable cooperation and patience.

Prague, March 2009

Chief editor:
Jiří Hřebíček

Main editors:
Jiří Hradec, Emil Pelikán, Ondřej Mirovský,
Werner Pillmann, Ivan Holoubek, Thomas Bandholz..

Contents

III. Workshop Papers

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

One Information Space, Billions of Sensors

Alessandro Annoni and Massimo Craglia
European Commission, Joint Research Centre
alessandro.annoni@jrc.it; massimo.craglia@jrc.it

Abstract: Interactions and inter-dependencies between governmental and social infrastructures become more and more complex and their governance requires new ways of thinking and new approaches. To face challenges related to the collection and availability of environmental information or information to support the protection of the environment at different levels, local, regional, global, a new approach is needed. We should be prepared to manage the heterogeneity of emerging dynamic systems and be able to transform increasing amount of real time available data into reliable, accurate, timely, and openly accessible information at the relevant geographic and temporal scales. This paper introduces some research challenges for the development of an improved, comprehensive and efficient eGovernment framework.

***Keywords**:* environment; spatial information infrastructures; digital earth; voluntary information.

1. INTRODUCTION

The importance of a geographic perspective to underpin public policy, business, and social interactions has gained considerable recognition during the last decade. Many governments world wide have been investing in internet-based infrastructures to support the sharing and effective use of spatial information and data, and a global effort to connect systems observing the earth is underway since 2003 under the aegis of the Group for Earth Observation (GEO). In parallel to these initiatives, private sector led developments in geo-browsers (e.g. Google Maps, Google Earth, and Microsoft Virtual Earth) have brought geography to hundreds of millions of users as a way to explore the Earth, find and visualise information, and become active participants and producers of data. The opportunities for increased interactions among citizens, governments, and businesses, and the development of a spatially-enabled Information Society have never seemed greater. The challenge however, is in channelling the many initiatives in a framework that is inclusive and reduces the digital divide; a framework that is open to new developments, without freezing the status quo, and supportive of increased dialogue and cooperation; one that leads from eGovernment to eGovernance i.e. from one-way interactions from government to citizens, to a participatory framework in which all stakeholders take part in developing a shared understanding, and in addressing key issues together. The Environment is a key test case for this new approach given the complexity of interactions and inter-dependencies among activities and decisions at multiple levels, local to global, that affect us all. With these considerations in mind, this paper reviews some of the key developments in the way we collect, organise, and share information about the planet (Section 2), puts forward a vision of the next generation Digital Earth (Section 3), and identifies some key research challenges we need to address to achieve that vision (Section 4), with recommendations for short term and mid-term action (Section 5).

2. KEY DEVELOPMENTS IN GEOGRAPHIC INFORMATION

It may be useful to group recent key developments into the four main streams discussed below.

2.1 Organising Geographic Information

This stream includes the many initiatives since the early 1990s aimed at increasing the availability and accessibility of geographic information through the development of spatial data infrastructures (SDIs). By the mid 1990s, Masser [1999] identified at least 11 SDIs at varying stages of development spanning large countries like the USA, Canada, and Australia, small ones like the Netherlands and Portugal, and developing nations like Malaysia, Indonesia, and Qatar. That first generation of SDIs was largely led by national mapping agencies and oriented towards the completion of national spatial databases addressing topography and other key layers of general use. The documentation of existing resources via metadata, and access mechanisms via catalogues and clearinghouses were other key features of these early developments.

Since then we have seen a rapid diffusion of SDIs world wide facilitated by the establishment in 1996 of the Global Spatial Data Infrastructures Association that has helped the promotion of best practice and sharing of experiences, and capacity building in the Americas, Africa, Asia and the Pacific and Europe. The nature of these more recent SDIs has also shifted with an increased number of stakeholder organizations engaged in the process. There is also a stronger emphasis on distributed data and processes, and the interoperability of services to discover, view, access, and integrate spatial information.

In Europe a major recent development has been the adoption of the INSPIRE Directive (2007/2/EC), establishing an infrastructure for spatial information in Europe to support Community environmental policies, and policies or activities which may have an impact on the environment (http://inspire.jrc.ec.europa.eu/). INSPIRE is based on the infrastructures for spatial information established and operated by the 27 Member States of the European Union, which are required to transpose the Directive by May 2009. The Directive addresses 34 spatial data themes needed for environmental applications, with key components specified through technical implementing rules covering:

1. Metadata to describe existing information resources so that they can be more easily found and accessed.
2. Harmonisation of key spatial data themes and related spatial services needed to support environmental policies in the Union.
3. Agreements on network services and technologies to allow discovery, viewing, transformation and download of information resources, as well as access to related services.
4. Policy agreements on sharing and access, and measures to monitor the implementation of the infrastructure.

With the adoption of the technical implementing rule for metadata, and those in progress concerning discovery and view services we are starting to have the key building blocks enabling an initial operating capacity in Europe to link the SDIs of the Member States and make existing geographic information more visible and accessible to users. Forthcoming specifications on the interoperability of spatial datasets and services, and sharing will then develop the infrastructure further to support environmental policy in Europe.

INSPIRE provides the legal framework and infrastructural layer upon which two other initiatives are being developed: (i) the Shared Environmental Information System (SEIS) is a collaborative initiative of the European Commission and the European Environment Agency to establish with the Member States an integrated and shared EU-wide environmental information system linking existing information flows related to EU environmental policies and legislation (http://ec.europa.eu/environment/seis/). (ii) the Global Monitoring for Environment and Security (GMES) initiative aims to provide operational information services based on Earth monitoring data from satellites and in-situ observations on water, air and land (http://www.gmes.info/). Whilst all three initiatives (INSPIRE, SEIS, GMES) pertain to the environmental domain, their complementarity can be portrayed as focusing on data, information, and services respectively.

2.2 Geography to Organise Information

This stream includes the development of geo-browsers (e.g. Google Earth, Microsoft Virtual Earth, NASA Worldwind, ESRI ArcGIS Explorer) which use the globe as mechanism to pan, zoom, and fly over the Earth's surface to areas of interest. Associated to these 3D representations of the Earth are also 2D applications (Google Maps, Microsoft Live Search Maps) that also allow users to add and share information via simple Application Programming Interfaces (APIs). These have enabled a turn around of the traditional approach in which government agencies (or the private sector) produce data, and citizens (or scientists) are the users, to one in which users become themselves producers (see also Section 2.3). Moreover, the widespread success of these platforms and applications, with hundreds of millions of users, have had the major benefit of popularizing geography among wide strata of the population, as well as enabling companies to add a spatial perspective to their business processes. Beyond visualization however, geobrowsers have become client applications for accessing a more complex infrastructure behind the scenes, based on a loosely connected information model and technical breakthroughs in horizontal (multiple machines) and vertical (multiple hardware/software tiers, together with data caching mechanisms). We could refer to this as the GeoWeb infrastructure to provide contrast to the traditional SDI approach.

2.3 Geosensors and systems

Geosensors can be defined as any device receiving and measuring environmental stimuli that can be geographically referenced. As such they include satellite-based sensors providing multi-spectral information about the Earth's surface (imagery, land cover, vegetation indices and so on), air-borne sensors for detailed imagery but also for laser scans of physical or man-made structures (LiDAR), and sensors near, on, or under the Earth's surface measuring anything from physical characteristics (pressure, temperature, humidity) and phenomena (wind, rain, earthquakes), to the tracking of animals, vehicles, and people. Large-scale networks of sensors have been in existence for several decades. What is novel is the web-enablement of these sensors and their networks so that individual sensors can be discovered, tasked, and accessed through web standards (sensor web), and that the networks can exchange information through interoperability arrangements.

It is worth noting that a new breed of sensor networks, called wireless sensor networks (WSN), have demonstrated the potential to revolutionize the way we acquire geospatial data. Different from the traditional, large-size, complex and costly sensor stations, a WSN typically consists of miniature, battery-powered nodes with power-efficient CPUs, short-range radios and low-cost sensors. The software that runs on the WSN nodes allows them to self-assemble into ad-hoc networks, such that the network can be easily deployed (e.g., sensors can be seeded from a low-flying airplane throughout hazardous areas) and data can be relayed across multiple hops and from long distances. This new class of sensing platform will provide an unprecedented volume of real-time geosensor data, along with high spatial and temporal resolution.

Linking existing systems and networks to achieve comprehensive, coordinated and sustained observations of the Earth system is the objective of the Global Earth Observation System of Systems (GEOSS). GEOSS is overseen by the Group on Earth Observations (GEO), an intergovernmental organization at the ministerial level comprising of 73 nations, the European Commission, and 52 international organizations. A major role of GEOSS is to promote scientific connections between the observation systems that constitute the system of systems, and promote applications across nine societal benefit areas addressing water, weather, ecosystems, agriculture, biodiversity, disasters, health, energy, and climate (http://earthobservations.org).

The GEOSS 10-Year Implementation Plan (2005-15) [GEO, 2005] underscores the point that the success of GEOSS will depend on data and information providers accepting and implementing a set of interoperability arrangements, including technical specifications for collecting, processing, storing, and disseminating shared data, metadata, and products. GEOSS interoperability will be based on non-proprietary standards, with preference to formal international standards. The architectural design of GEOSS mirrors to a large extent that of the SDIs discussed in Section 2.1, with emphasis on metadata, catalogues, clearinghouses and portals, and registries of existing standards across different communities.

2.4 People as Sensors

A special type of "sensors" is that of citizens volunteering geographic information. Goodchild [2007] argues that there is a long tradition of non specialists contributing to the collection of scientific information such as the case of the Christmas Bird Count (http://www.audubon.org/bird/cbc) or the collection of weather information in the GLOBE programme (http://www.globe.gov) but that only recently the convergence of greater access to broadband connections, the availability of Global Positioning Systems at affordable prices, and more participative forms of interaction on the Web (Web 2.0) have enabled vast numbers of individuals to create and share geographic information. Platforms such as Google Maps and Microsoft Live Search Maps have made it possible to publish and make geographically searchable user-generated content to an unprecedented rate. Initiatives such as Wikimapia (http://www.wikimapia.org) and OpenStreetMap (www.openstreetmap.org) show how organised volunteered information can challenge traditional data suppliers with good-quality products that are openly accessible to all. As observed by Goodchild, the potential of up to 6 billion human sensors to monitor the state of the environment, validate global models with local knowledge, and provide information that only humans can capture (e.g. emotions, and perceptions like fear of crime) is vast and has yet to be fully exploited. An excellent recent example of the opportunities is provided by the Brasilian TV network Globo, that opened in September 2008 an interactive website based on Google Maps to monitor the state of the Amazon forest, and provide the opportunity to citizens to report instances of illegal logging and forest clearances. Whilst having a "modest" target of 1 million users in the first year, the site (www.globoamazonia.com) already received 41 million reports by individuals in the three months to December 2008, making it one of the most successful web sites in Brazil, and obtaining real impact through follow up legal initiatives and Parliamentary enquiries.

Against developments such as the ones described above, current SDIs, and GEOSS, do to seem well equipped neither to handle rapidly changing data, nor to consider the major contribution of the public as producers of relevant geographic and environmental data. Aside from technological considerations, there is also a cultural barrier in the perception that only "official" information i.e. from certified government sources is reliable, while citizens' information is not. Of course issues of validity, reliability, and trust are crucial, but it is also worth noting here that for example, that during the forest fires in California in November 2008, in which 230 houses were burned, images and voluntary reports uploaded within minutes on Flickr were a key vehicle to alert people of incoming danger and support evacuation, while official information was way too slow, and even Google could not index the relevant information fast enough [Goodchild 2009].

3. TOWARDS THE NEXT-GENERATION DIGITAL EARTH

Ten years ago, U.S. Vice-President Al Gore articulated a vision of "Digital Earth" as a multi-resolution, three-dimensional representation of the planet that would make it possible to find, visualize, and make sense of vast amounts of geo-referenced information on the physical and social environment. Such a system would allow users to navigate through space and time, access to historical data as well as future predictions based for example on environmental models, and support access and use by scientists, policy-makers, and children alike [Gore 1999]. As indicated in the earlier sections of this paper, significant progress is being made in the availability and quality of environmental and geographic information at our disposal, and in connecting information systems and new sources of data. The diffusion of SDIs, the efforts towards GEOSS, and developments in industry and civil society have major contributions to make in helping us understand how the Earth's environment is changing, and what are the consequences for human civilization, which are among the fundamental questions of our time. Taken individually however, none of these developments can achieve the objective or get us to the vision of Digital Earth put forward more than 10 years ago. Now is the time to set a new vision of what can be achieved within the next 5-10 years, building on what is already existing, bridging the gaps, and overcoming current limitations. Below are the initial elements of such a vision.

1. *Not one Digital Earth, but multiple connected globes/infrastructures addressing the needs of different audiences: citizens, communities, policy-makers, scientists, educationalists.*
   Each audience has a distinct set of needs for information about Earth and its future, so we anticipate that each would be accommodated by a specially designed Digital Earth. One might encourage members of the general public to contribute their own observations, while another would present only the most rigorously obtained scientific results. Each would be a view, however, of a single coordinated, distributed data resource. Different views would require different levels of detail, of geometry, imagery, attributes (semantics), etc.

2. *Problem oriented: e.g. environment, health, societal benefit areas, and transparent on the impacts of technologies on the environment.*
   While it is important that specific problems be addressed in focused ways, Digital Earth should still clarify the interactions between problems and objectives – the difficulties of achieving one objective, such as reducing the costs of energy, with others such as impacts on the environment and food production.

3. *Allowing search through time and space to find similar/analogue situations with real time data from both sensors and humans (different from what existing GIS can do, and different from adding analytical functions to a virtual globe).*
   While there are strong affinities between the current generation of virtual globes and earlier GIS technology, it is clear that users expect virtual globes to answer a different kind of query, one that is less precise and quantitative, and more attuned to exploration and browsing. For example, one popular use of Google Earth is to search the globe for similar conditions – for example, to find areas that are as vulnerable to tsunamis as the coasts of the Indian Ocean.

4. *Asking questions about change, identification of anomalies in space in both human and environmental domains (flag things that are not consistent with their surroundings in real time).*
   One of the most compelling benefits of satellite images, maps, and virtual globes lies in their ability to provide context, by displaying information in its correct geographic position. The next generation of Digital Earth should allow for rapid search for geographic anomalies, that is, situations that are inconsistent with their geographic context, such as outbreaks of disease, biodiversity hotspots, or anomalous levels of air pollution.

5. *Enabling access to data, information, services, and models as well as scenarios and forecasts: from simple queries to complex analyses across the environmental and social domains.*
   One of the challenging issues today is to combine environmental modelling and forecasting with its socio-economic impacts. Traditional flood forecasting or mapping of natural hazards risk zones loose their value if their social impact is not assessed. Meanwhile such models have immediate economic consequences (e.g. property value reduction in the case of identified risk) and for this reasons both model reliability and appropriate communication tools (including visualization) are strongly required.

6. *Supporting the visualization of abstract concepts and data types (e.g. low income, poor health, and semantics).*
   Advances in dynamic visualization environments (see, for example, www.gapminder.org) show strong potential for decision support and increased understanding of global, complex, and abstract phenomena. Bringing these capabilities to the next generation Digital Earth will turn these into important tools for education, awareness-raising, and informed decision making. Different perspectives on phenomena like poverty or health and their indicators can now be made explicit through ontologies, and mappings between them have become possible.

7. *Based on open access, and participation across multiple technological platforms, and media (e.g. text, voice and multi-media).*
   The geoinfo community may have a great deal to learn from the wider multimedia community. For this to happen, the emphasis on maps (fixed geometry plus labels) as central object of inquiry, study and navigation, must shift and dynamic elements must be incorporated. There is promising work already underway that integrates georeferenced moving video with other static geographic data sets. Open access policies are necessary to foster participation and inclusion across communities.

8. *Engaging, interactive, exploratory, and a laboratory for learning and for multi-disciplinary education and science.*
   The notion of virtual collaboratories is a key feature of e-Science to support the multidisciplinary exchange of knowledge across scientific teams dispersed at multiple locations. In other fields interactive learning tools, distance learning, but also location-based games offer platforms and lessons that can be built upon to develop teaching, learning, and sharing environments for multiple audiences that must be fun to use as well as scientifically rigorous.

4. RESEARCH NEEDED TO ACHIEVE THE VISION

1. *Information integration (multi source and heterogeneous, multi-disciplinary, multi-temporal, multi-resolution, and multi-media, multi-lingual).*
   Despite substantial progress, our ability to integrate geographic information from multiple sources is still quite limited. We need a better understanding of the statistical problems of integration across scales, the linguistic problems of integrating across languages, and the semantic problems of integrating across disciplines. This will require a substantial effort by a number of collaborating disciplines: computer science, information science, and the domain disciplines.

2. *Space-time analysis and modelling (i.e. universal elements and language for dynamic modelling, algebra of space-time change).*
   The next generation of Digital Earth should provide a powerful platform for simulating the human and physical processes that operate on the Earth's surface. While such models have been developed in many domains, they use a myriad of approaches that are impossible to couple or integrate. Fundamental research is needed to develop a comprehensive language for simulation, and the software components needed to make simulations easily interoperable across disciplinary boundaries.

3. *Schemes for tiling the curved surface of the Earth and for use in data management, analysis, simulation, visualization.*
   Each of the current generation of virtual globes uses its own approach to structuring data. These systems are optimized for storage and display purposes, but have limitations for the analysis of global scale data and processes. Research on optimal structuring and indexing schemes has been under way for the past two decades, but we do not yet know how to design an optimal scheme that can support massive simulation of Earth-surface processes.

4. *Intelligent descriptions (automatic, user driven) of data, services, processes, models, searching and filtering.*
   Good progress has been made on standardizing the description of data, through metadata and syntax standards. Still missing, however, are adequate standards for the description of models, processes, and services that can support search and the assessment of fitness for use. Also user feedback on fitness-for-purpose will help us develop much more usable and useful metadata (Metadata 2.0 to build on Web 2.0).

5. *Visualization of abstract concepts in space.*
   Transformations from lower level, physical observations, to abstract concepts like quality of life or vulnerability need to be modelled and implemented, to support visualization and reasoning. The wealth of data sources that can now be tied into Digital

Earth, together with advances in complex system modelling and semantic mappings, provide a richly equipped laboratory for domain and information scientists to enable new uses of Digital Earth. Progressing from specific, well-defined case studies on well-defined scientific phenomena to more complex cases of socially defined and negotiated notions promises gradual, but significant progress.

6. *Computational infrastructures to implement vision (architecture, data structures, indexing, interfaces).*
   Advances in technology, information structuring, and organization of the IT infrastructure have already made it possible to come close to the vision of Digital Earth. The interoperability of multiple systems delivering data, information, and models in real-time from multiple sources, including passive sensors and humans, can increase by several orders of magnitude the quantity of information posing new challenges to deliver reliable, and timely quality information to multiple and diverse audiences through multiple access media including voice.

7. *Trust, reputation and quality models for contributed information and services.*
   Progress from traditional provider perspectives on data quality to broader notions of fitness for use, trust, and reputation is already happening in the context of Volunteered Geographic Information but not yet in government-led infrastructures. There is a need to develop metrics of trustworthiness linked to context, applications, and response time and then be able to encode these metrics into software so that they can support machine-too-machine interactions.

8. *Governance models and collaborative frameworks (business, institutional, voluntary, communities of practice).*
   The emergence of hybrid infrastructures can already be observed combining both voluntary and institutional data (e.g. using the Google platform). Without a mechanism to clearly distinguish the different nature of the data, users will be reluctant to take any formal decision (still recognising the value of non-institutional data). An important challenge is to build new models for reciprocal validation of data made available through collaborative frameworks (e.g. validation of precise quantitative comprehensive information about air quality collected through few professional measuring stations to be combined with air quality data collected through mobiles equipped with appropriate nanotechnologies measuring only a limited number of parameters). Communities of practice should be better engaged in Governmental decisions sharing their data and knowledge in an easier and effective manner

9. *Data sharing and open access policies.*
   More systematic and comparable evidence is needed on the impact of different access policies on organizations and society in the face of multiple pressures on government budgets and increasing inter-organizational competition for funding. If such evidence was already in place, we would not still be facing tensions surrounding data access and sharing policies. More work is also needed on incentives and barriers to data sharing at individual, inter, and intra organizational levels to support participation and sustainability.

10. *Social and economic impacts of Digital Earth*.
    Appropriate theoretical and methodological frameworks to assess the social and economic impacts of geo-spatial information, and related infrastructures are still poorly developed but are urgently needed to justify the initial investment and the long-term sustainability of the infrastructure. The costs of not acting also need attention (in a similar vein as the Stern [2006] Review on Climate Change) at different scales from global, to regional, national, and local to inform appropriate funding and business models.

5. CONCLUSIONS

The global financial and economic crisis we are experiencing today requires coordinated action and leadership. Investment in research, science, and technology can contribute to set the bases for a recovery that is sustainable and environmentally conscious. We have outlined in this paper a vision for the next generation Digital Earth to support a greater understanding of environmental change and its consequences on society. We have also identified some key research topics that can be addressed in the next 3-5 years through a coordinated international effort. A priority however, is increasing the dialogue and collaboration between government-led and citizens-led initiatives; between "official" information infrastructures for policy and science, and social networking at the grass-routes. It is inconceivable to build an earth observing system without the contribution of the billons of people who inhabit it. For this we need to leave behind the ivory towers and work together to build the bridges between communities and groups of interest so that each can add value for the benefit of all. The technologies to do so are already available. The will to act is the immediate priority.

ACKNOWLEDGEMENTS

This contribution builds on the position paper on Next Generation Digital Earth [Craglia et al., 2008] that came from the specialist meeting on Virtual Globes organised in June 2008 by the Vespucci Initiative for the Advancement of Geographic Information Science (www.vespucci.org) and the Joint Research Centre of the European Commission. We are grateful to Michael F. Goodchild, Gilberto Camara, Michael Gould, Werner Kuhn, David Mark, Ian Masser, David Maguire, Steve Liang, and Ed Parsons who helped develop that paper and vision. To develop the vision further a Forum has been set up following the main themes of that paper, and we would welcome your contributions opinions. Please see http://forums.ec.europa.eu/nextgenerationdigitalearth/

REFERENCES

Craglia, M., Goodchild M.F., et al. Next-Generation Digital Earth: A position paper from the Vespucci Initiative for the Advancement of Geographic Information Science, *International Journal of Spatial Data Infrastructures Research, 3*, 146-167, 2008. http://ijsdir.jrc.ec.europa.eu/index.php/ijsdir/article/view/119/99

Goodchild M.F., Citizens as voluntary sensors: spatial data infrastructure in the world of Web 2.0. *International Journal of Spatial Data Infrastructures Research, 2*, 24–32. 2007. http://ijsdir.jrc.ec.europa.eu/index.php/ijsdir/article/view/28/22

Goodchild M.F., *SDI, the Citizens, and Web 2.0*. Presentation at the II Trobada de la IDEC, Barcelona, 2-3 March. 2009.

Gore A. The Digital Earth: Understanding our planet in the 21st Century. *Photogrammetric Engineering and Remote Sensing, 65* (5), 528, 1999.

Group on Earth Observation (GEO) Secretariat (Ed.), *GEOSS 10-Year Implementation Plan*, GEO, Geneva, 2005. http://earthobservations.org/documents/10-Year%20Implementation%20Plan.pdf

Masser I., All shapes and sizes: the first generation of national spatial data infrastructures, *International Journal of Geographical Information Science, 13*, 67-84. 1999

Stern N. *Stern Review on the Economics of Climate Change*. Cambridge University Press. 2006. http://www.hm-treasury.gov.uk/independent_reviews/stern_review_economics_climate_change/sternreview_index.cfm

Session 1

ICT research towards a Single Information Space in Europe for the Environment

Organized by **Jiří Hřebíček**, Colette Maloney, Manuel Monteiro, Andrea Rizzoli and Alessandro Annoni

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

GENESIS, a flexible solution contributing to the SISE deployment

Christian ALEGRE, Stéphane BONNOT
Thales Alenia Space, Optical Observation and Science Business Unit
(christian.alegre@thalesaleniaspace.com)
(stephane.bonnot@thalesaleniaspace.com)

Abstract:
The need for a Single Information Space for Environment in Europe (SISE concept), an EU wide electronic space avoiding technical or regulatory blockages and facilitating interoperability between information systems and collaboration between actors, is highlighted by DG INFSO in FP7 ICT work program. It is also underlying in other European directives or initiatives like INSPIRE and SEIS.

GENESIS Integrated Project, funded by DG INFSO in the frame of the FP7 ICT work programme, responds to this need, by proposing a reference European solution and by demonstrating the efficiency of the solution through thematic pilots dealing with the environment management and its impact on health.

GENESIS project is lead in synergy with various other EC or ESA projects also focused on interoperability concepts and standards.

GENESIS has started with the analysis of the state of the art in the considered thematic domains (Air Quality, Fresh Water Quality, Coastal Water Quality) as well as in the ICT technology domain then by the identification of the main gaps and potential improvements.

***Keywords**:* SISE ; environment management ; standard scalable solution ; services oriented architecture ; Air quality ; Water Quality.

1. INTRODUCTION

1.1 Context

The environmental management is one of the main target for which DG ENV has launched the INSPIRE initiative aimed to make harmonized and high-quality geographic information readily available for Communities and for the citizen to access spatial information.

The need for a Single Information Space for Environment, an EU wide electronic space allowing interoperability and avoiding technical or regulatory blockages, is highlighted by DG INFSO in FP7 ICT work program.

The Single Information Space for Environment in Europe (SISE, see figure 1) is expected to facilitate the access by all actors (decision makers, scientists, users involved in environment management processes or activities, citizens) to shared qualified information.

GENESIS responds to both DGs' objectives by proposing a strategic integrated approach leading to a reference European solution and by demonstrating the efficiency of the solution through thematic pilots dealing with the environment management and its impact on health.

The reference solution proposed by GENESIS is open, standard and flexible, it can be easily instantiated for various thematic fields and for various contexts all over Europe.

The solution can be deployed first on a set of pilots then on a wider scale to facilitate the overall European environment monitoring and control. Further on it can leverage the market of added value services (based on simpler publication of data and processing, easier combination or fusion of various data, transparent integration of Web technologies, …).

GENESIS project is supported by the World Health Organisation, who has pointed out the interest of the combination of information coming from various data sources, including monitoring, satellite assessments and modelling, to improve ability for more precise exposure assessment, and the interest of the information on population exposure on a particulate matter, generated by GENESIS, as a potential input of the Environment and Health ENHIS information system.

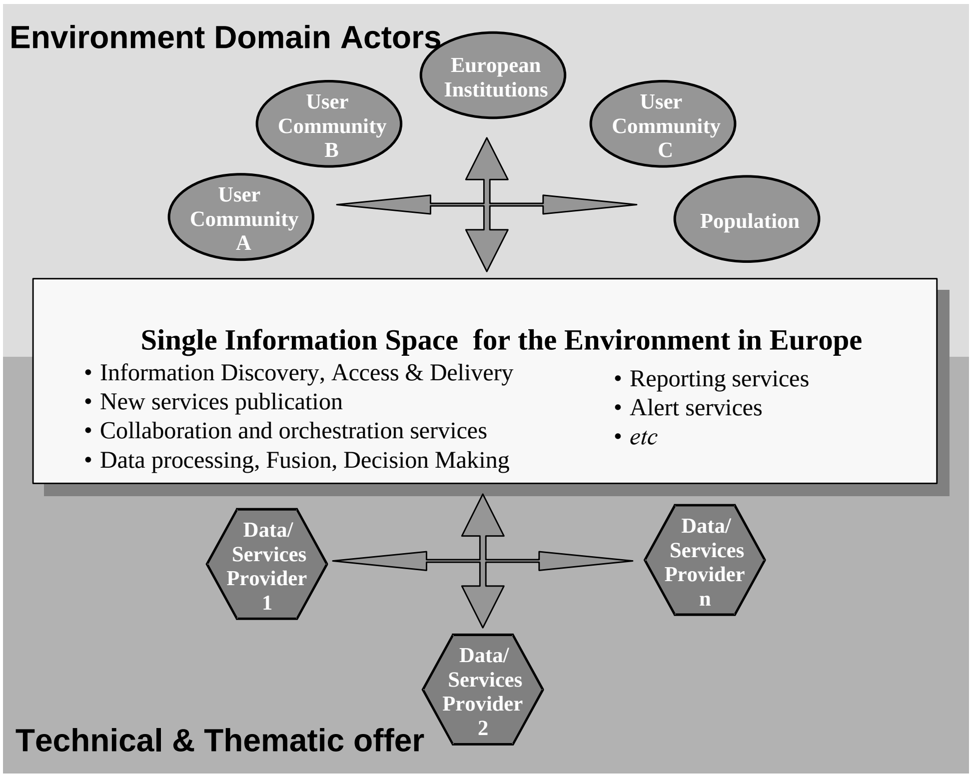


Figure 1 : SISE concept

1. 2 GENESIS objectives

The global target of GENESIS project is to contribute to the definition and the set-up of the Single Information Space for the Environment in Europe, for the sake of various thematic communities in relation with environment management, in particular health community.

In other words, GENESIS must provide the thematic communities or users, who are not necessarily aware or expert of new technologies, with an easier access to these technologies.

More precisely , the following scientific objectives can be identified :

- The selection and promotion of new technologies helping to support the SISE concept,

- The definition and development of a generic (thematic neutral) solution, implemented by the GENESIS software, optimal both in term of use of new technologies and in term of coverage of thematic needs,
- The support of the thematic actors involved in the GENESIS pilots for the use of the technologies and of the generic solution proposed by GENESIS.

The objectives can be further breakdown into several sub-objectives :

- Identification of the domain-neutral requirements for the solution, applicable to generic collaborative information systems for Environment management and analysis of environmental impact on Health,
- Identification of the available technologies that can be used to support these requirements,
- Definition, design and implementation of a generic solution supported by the GENESIS software covering the previous domain-neutral requirements, with state of the art and advanced technologies
- Demonstration and validation of the solution through pilots in several thematic domains targeted by the project (Air Quality, Fresh surface water quality, Coastal bathing water quality).
- Adjustment of the solution taking into account the feedback from use of the preliminary version of the solution :
    - ➢ The thematic partners build their thematic pilots by instantiating the GENESIS software with integration/connection of the relevant specific thematic components,
    - ➢ The thematic partners evaluate the quality of the GENESIS solution together with representative users of the domain including various health users.

2. QUICK LOOK ON THE SOLUTION

The definition of the generic solution is in progress, it is derived from the convergence of two processes :

- the analysis of the state of the art in the considered thematic domains and the identification of the main gaps and possible improvements,
- the analysis of relevant ICT technologies (like multi-linguality and ontology, sensor networks architecture, near-real-time data fusion, dynamic work flows, customizable portal,….) that can provide an effective advantage to the environment management actors.

The two parallel processes of analysis are performed in synergy with major European or worldwide initiatives or projects :

- compliance with INSPIRE directive is “the first driver” (and INSPIRE intermediate results are used whenever available),
- the architecture definition and ICT technologies assessment builds on :
    - ➢ the recommendations coming from INSPIRE, GEOSS, …
    - ➢ the recommendations or results coming from the major European projects like FP7 GIGAS, FP6 WIN, ORCHESTRA, SANY, OSIRIS,..., ESA MASS-SSE project, or GMES projects like LIMES, PROMOTE, GEMS, HALO, BOSS4GMES.

The solution is fully based on standards and harmonization results. There are two distinct technological sub-projects that are in charge of :

- the assessment of advanced technologies and the preparation of their deployment for the benefit of the thematic communities,
- the definition and implementation of an “information system set-up framework”, which can be used by the thematic communities to build or upgrade information systems for environment management with high interoperability.

The “information system set-up framework” is :

- based on a Service Oriented Architecture, compatible with INSPIRE and GEOSS architecture guidelines,
- open and sustainable as based on de facto and emerging standards (OGC, OASIS, INSPIRE,…).
- implemented as a software package made as much as possible open source software or components.

The software package is composed of generic (thematic-neutral) software components (Web services, Web portal components,..), data & metadata models and toolkits (for the configuration and the deployment of the information systems in specific thematic contexts), accompanied by the relevant documentation.
For example Web portal components include a generic Web portal plus a Web portal framework with modular components (portlets) that can be used to build customised thematic Web Portals.

The core set of Service implementations includes :

- Discovery services (for services, sensors, EO datasets and GIS datasets);
- Data access services (for EO data, in-situ data, GIS data and non-geo data);
- Geo-information and decision support services (including data fusion services, geo-information services, service orchestration services, decision support services and communication services);
- User management services (including services for authentication, authorization and user registration);
- Human interaction services.

The toolkits help to instantiate the software for a specific domain (e.g. to create thematic workflows, to connect legacy systems, to compose thematic web portals, to deploy and administrate the resulting information system …).

3. PROJECT OVERVIEW

3. 1 Project team

The GENESIS Consortium brings together a wide panel of multidisciplinary skills and experiences in the major European projects and initiatives, in both ICT and thematic application domains :

- participants having skills and experiences in environmental domains as well as in health issues are key actors for the identification of users requirements, definition and implementation of the pilots, evaluation and take-up of the solution;
- participants having skills and experiences in ICT and related domains are key actors to define a reference collaborative architecture, and to make an efficient and sustainable solution.

IT expertises, which are required for defining and developing the GENESIS solution, are strongly represented in the Consortium, for example in the following domains :

- Sensors networks ,
- Data-metadata modelling,
- Data fusion ,
- Ontology and Multi-linguality,
- Geographical Information Systems,
- Workflows for service chaining,
- …

Beyond IT expertises, strong competences are available in the GENESIS team in other domains like :

- The Environment and Health issues,
- The GMES core services in Air Quality and Coastal Water quality domains,
- The GEO/GEOSS programme,
- The main European or International standardization and harmonization initiatives (INSPIRE, ISO, OGC, …),
- The development of the ESA Services Support Environment.

3. 2 Project work plan

The rationale of the project is the following :

- The thematic actors analyse the state of the art in the considered thematic domains, and identify the main gaps and possible improvements in relation with the generic IT solution.
- The technological actors, who are "providers" of the GENESIS solution, assess the available advanced technology, in relation with the needs and gaps expressed by the Thematic actors, select the best new technologies to be included in the GENESIS solution, with respect to their pertinence and their maturity level, then perform all activities up to the release of the generic IT solution,
- Several thematic pilots allow to use and evaluate the generic IT solution,
- The thematic actors provide feedbacks so that the solution can be refined and completed.

To fit with this rationale, the project includes several technical sub-projects :

- Thematic sub projects that participate all along the GENESIS project to model the environmental and health domains, to express related requirements and needs in terms of IT system, to evaluate the GENESIS solution in a real implementation integrating their usual thematic services and finally to valid the GENESIS interoperability and modularity through various pilot experimentations.
- Technological sub-projects in charge of new technology assessment and new technology integration into the generic solution.
- A Cross cutting sub-project that addresses the definition of the generic solution by finding the best convergence between the IT needs expressed by the thematic teams and the effective IT technology that can be put aboard a sustainable easy to use/easy to set-up solution.

The pilots focus on thematic fields of Air Quality, Fresh Water Quality, Coastal Water Quality and their impact on Health.

The development process is based on a strong and early synergy between thematic and technological activities ensuring the capitalization of users' feedback : the thematic actors, who are "clients" of the solution, play a leading role in the selection of the innovative technologies to be integrated to the final solution, then in the validation of the solution in real applicative context.
Furthermore a staged development process is used to adjust the final solution taking into account the feedback received from the users ; two successive iterations are performed :

- Iteration 1 is based on the early definition, implementation and use of a preliminary solution,
- Iteration 2 is based on the final solution, which takes into account the feedback from the effective use of the preliminary solution .

This incremental process has some other advantages :

- It facilitates the integration within the project (in the final release) of various technologies which are not fully mature in the early stage of the project,

- It facilitates the synchronisation with various initiatives in progress in parallel of GENESIS project (INSPIRE, GEOSS,…),
- It helps to disseminate elements of the solution, to train and support the various users or stakeholders, in order to get a progressive take-up of the solution and the build of a medium term deployment plan.

3. 3 Project status

GENESIS has started by the analysis of the state of the art in the considered thematic domains as well as in the ICT technology domain then by the identification of the main gaps and possible improvements in relation with GENESIS solution.

After six months, one internal milestone has been passed and the second is in progress :

- Passed milestone : User Requirement Review 1 corresponding to a convergence on a first set of Generic User Requirements (thematic neutral requirements extracted from the Thematic User Requirements expressed the different pilots)
- In-progress milestone : Preliminary Design Review 1 corresponding to a convergence on a preliminary architecture and list of Generic Services Requirements.

Contributions to GEOSS or OGC work group are also planned or in progress.

4. CONCLUSION

The main benefits of the GENESIS solution is to facilitate the set-up of fully interoperable information systems with the following capabilities :

- Easy access by different actors to environmental data and related data at different levels, thanks to INSPIRE harmonisation,
- Improved near real time environment monitoring,
- Improved decision making based on near real time environmental parameters and fusion of various sources and types of data,
- Improved alert and response processes,
- Enhanced visualization , combination and correlation of static or near-real-time information,
- Easier consolidation at European level, in particular by the support of recurrent activities by automated work flows,
- Progressive constitution of a set of reference data models in Europe to support the various thematic domains.

The GENESIS project represents an important step in the operational environmental management in Europe. It paves the way to an effective wide deployment of the solution in various contexts :

- the progressive set-up of the Single Information Space for Environment in Europe, in line with INSPIRE directives
- the implementation at member state level of European environmental frameworks.

Its flexible generic solution can also contribute to :

- GMES core and downstream services framework ,
- The implementation of the SEIS principles,
- The support of GEO/GEOSS future architecture by generic building blocks.

ACKNOWLEDGEMENTS

The authors wish to thank the European Commission for the financial support of the project.

The list of GENESIS partners follows:

- Thales Alenia Space France (Fr)
- Deutsches Centrum fuer Luft und Raumfahrt E.V (DLR) (De)
- Sogreah (Fr)
- ACRI-ST (Fr)
- Spacebel (Be)
- University of Munster (WWU) (De)
- EC-JRC-IES & EC-JRC-IHCP (Eu)
- 4CT Technologies (Be)
- GMV Aerospace and Defense (Sp)
- GIM (Be)
- INTECS (It)
- EBM Web Sourcing (Fr)
- ERDAS (Be)
- British Publisher Ltd (UK)
- Cambridge Environmental Research Consultants Ltd (UK)
- Imperial College of Science Technology and Medecine (UK)
- CHU Nice (Fr)
- Ludwig Maximilians University München (De)
- Institute fuer Ostseeforschung Warnemuende an der Universitaet Rostock (De)
- Institute Meteorologii i Gospordarki Wodnej (Pl)
- Commisariat à l'Energie Atomique (Fr)
- Ingenieurgesellschaft Prof. Kobus Und Partner GMBH (De)
- Institut Pasteur (Fr)
- Kaipedos Universitetas (Lt)
- Fraunhofer Institute (De)
- ORACLE (Be)
- University of West Scotland (UK)
- Bureau de Recherches Géologiques et Minières (Fr)
- Research Studio Austria (At)

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Building the SISE: an environmental ontology

Pilar León Araúz[a], Pedro J. Magaña[b] and Pamela Faber[a]

[a] *Department of Translation and Interpreting, University of Granada, Granada, Spain (pleon &pfaber@ugr.es)*

[b] *Department of Computer Science and Artificial Intelligence, University of Granada, Granada, Spain (pmagana@decsai.ugr.es)*

Abstract: One of the main goals of the Single Information Space in Europe for the Environment (SISE) project is to provide some sort of integrated environmental information space. This task has much in common with the Semantic Web initiative. In this paper, we present part of the conceptual structure of wastewater treatment within EcoLexicon. Consistent with Semantic Web design guidelines, our domain knowledge needs to be represented by ontology modelling in order to provide a suitable schema for sharing and reusing semantic resources. Furthermore, reasoning techniques can be applied to discover and extract new information, thus increasing the richness of potential queries. Nevertheless, many projects have not been conceived as such from the beginning. As a result, these legacy systems, as it is the case of EcoLexicon, need to find some mechanism to get integrated in ontological models. On the other hand, terminological resources can find in ontologies a powerful representational model. EcoLexicon is structured around an Environmental Event (EE) which provides the conceptual underpinnings for the location of conceptual sub-hierarchies. Based on the ontology semantics, sub-hierarchies can be visualized in the form of a user-friendly conceptual network. Concepts are then linked to each other through an inventory of relations, which largely depend on the type of entity being described, its nature, and its own relational power. They can be enhanced by an additional degree of OWL semantic expressiveness. However, concepts tend to be intermingled in different situations, and they are not always related to the same concepts or through the same relations. As a result, at a more specific level, concepts appear in more fine-grained representations, where context-dependent dimensions are added, restricted or highlighted. The next step in our project entails extending the procedure followed in the wastewater treatment context to other domains of the relational database.

***Keywords*:** Environmental data; Environmental information; Environmental knowledge; Environmental information space; Conceptual structure; Ontology.

1. INTRODUCTION

One of the main goals of the Single Information Space in Europe for the Environment (SISE) project is to "provide some sort of integrated information space in which environmental data and information will be combined with knowledge for a decision support of environmental protection and sustainable development", as stated by Hřebíček et al. [2007]. To accomplish this objective, it is necessary to provide a knowledge framework capable of managing and integrating information from different sources. This task has much in common with the Semantic Web initiative, from which the e-environment community can benefit, especially from previously developed technologies that can provide a foundation for other initiatives.

In this sense, the e-environment resource presented in this paper (EcoLexicon) contributes to the development of the SISE, from both a linguistic and knowledge representation perspective. Our approach enhances knowledge exchange and offers easy access to the conceptual structures underlying the environmental domain. Moreover, it facilitates learning and communication, and eliminates conceptual and terminological confusion. These objectives also help to raise public awareness of environmental issues and contribute to the standardization of conceptual designations in different languages, something that also promotes a shared knowledge at an international level.

In this paper, we present part of the conceptual structure of WASTEWATER TREATMENT. This is a crucial sub-domain because it is linked to many specialized subject fields, such as chemistry, environmental engineering, microbiology, etc. Different disciplines deal with the same subject in different terms and, consequently, meaning can vary. This signifies that terminological standardization is imperative.

2. ONTOLOGIES AND TERMINOLOGICAL KNOWLEDGE BASES

Consistent with Semantic Web design guidelines, our domain knowledge needs to be represented by ontology modelling. An ontology defines the basic terms and relations comprising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary. Consequently, ontologies provide a suitable schema for sharing and reusing semantic resources. Furthermore, reasoning techniques can be applied to discover and extract new information, thus increasing the richness of potential queries.

Nevertheless, many projects have not been conceived as such from the beginning. As a result, these legacy systems need to find some mechanism to get integrated in ontological models. This issue has been widely studied in the Semantic Web community. However, it is a problem that has not yet been resolved.

Most such information comes from relational databases (RDB), as it is the case of EcoLexicon. In our approach, we emphasize the importance of storing semantic information in the ontology, while leaving the rest in the relational database. In this way, we can continue using the new ontological system, while at the same time feeding the database. This entails linking RDB stored information with an ontological system. This procedure can be useful for projects with similar purposes, especially during ontology mediation.

Nevertheless, this is not an easy task, since both representational models have remarkable differences. In contrast to relational databases, ontologies are highly expressive relational structures where concepts are described in very similar terms to those used by humans. Thus, relational models are suited to organize data structure and integrity, whereas ontologies try to specify the meaning of their underlying conceptualization [Barrasa, 2007].

In this sense, terminological resources can find in ontologies a powerful representational model. In turn, the design of ontologies can benefit from the theoretical background of linguistics, especially from cognitive approaches. Meyer et al. [1992] were one of the first terminologists to perceive that term bases would be more useful if their organization bore some resemblance to the way concepts are represented in the mind. As a result, what used to be simple repositories of terms and concepts became terminological knowledge bases (TKB), since they began to constitute authentic information systems that highlight conceptual interrelatedness and multidimensionality [Bowker, 1997].

In the same vein, according to Grinev and Klepalchenko [1999], the description of specialized domains can in many cases be based on the events that generally take place in them. This is the reason why our TKB is structured around an Environmental Event (EE) which provides the conceptual underpinnings for the location of conceptual sub-hierarchies [Faber et al. 2006].

The EE is based on the cognitive linguistics view of frames and semantic roles. According to Fillmore and Atkins [1992], frames are defined as a network of concepts related in such a way that one concept evokes the entire system. In contrast, in terms of specialized language applications Faber et al. [2007] describe them as dynamic structures that can streamline the action-environment interface of knowledge.

Consequently, the upper-level classes in our ontology correspond to the basic semantic roles described in the EE (AGENT-PROCESS-PATIENT-RESULT). And, as shown in Figure 1, all classes constitute a general knowledge hierarchy derived from each of them. This structure enables users to gain a better understanding of the complexity of environmental events, since they give a process-oriented general overview of the domain:

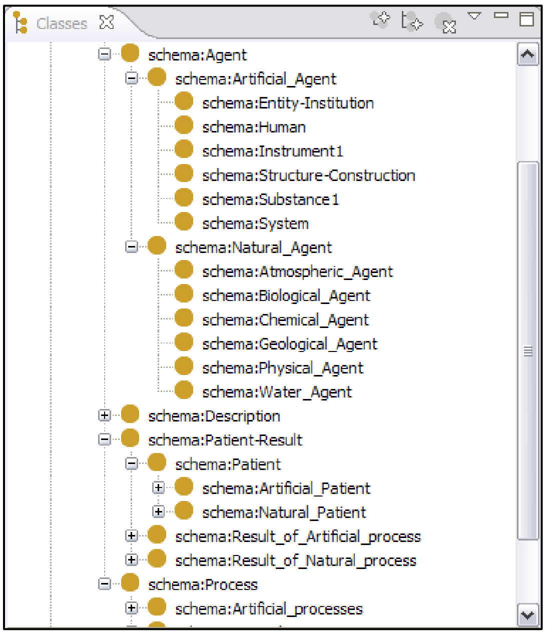


Figure 1. Ontological classes

These ontological classes are fed through the extraction of stored information in the database. This is done by using the D2RQ tool, which provides a usage scenario where relational databases are maintained as non-legacy applications [Bizer and Seaborne, 2004]. D2RQ is a declarative language to describe mappings between both systems. Moreover, these mappings can be conditional, which allows for feeding every class just with its corresponding instances.

3. CONCEPTS AND CONTEXT

Based on the ontology semantics, instances can be visualized in the form of a hierarchical user-friendly conceptual network. Concepts are linked to each other through an inventory of both vertical and horizontal relations, some of which are domain-specific and show their own internal hierarchical structure.

However, the determination of non-taxonomic conceptual relationships is not well-researched, and there is no consensus regarding what type of conceptual relationships should be modelled in a particular ontology. In our experience, relations largely depend on the type of entity being described, its nature, and its own relational power [Faber et al., 2008].

Thus, in EcoLexicon, the combinatorial potential of relations is based on the following criteria, some of which will be exemplified with concepts related to the WASTEWATER TREATMENT context:

- *is_a*: this generic-specific relation reflects hierarchical inheritance in the conceptual network of the domain. All entities and events are categorized building a hierarchical chain that is directly linked to the semantic roles of the event structure. Thus, any concept is linked to its immediate superordinate concept or to several concepts in cases of multidimensionality. For example, MAIN TRUNK SEWER → SEWER PIPE → PIPE → INSTRUMENT → AGENT (semantic role).

- *part_of*: this relation also reflects the hierarchical structure of the domain. In the case of physical objects, this relation directly refers to the parts of each concept. In the case of mental objects or processes, this reflection generally refers to phases. In the same way that objects are incomplete and can even lose their identity without one or more of their constituent parts, processes are incomplete without one or more of their phases (e.g. SCREENING *part_of* PRELIMINARY TREATMENT).

- *made_of*: this relation links both artificial and natural objects to the material they are made of, and thus bears a certain resemblance to the *part-of* relation without being the same. Even though the material of an object is part of it, this relation is different from the *part_of* relation since sometimes material is variable and it is not a distinguishable physical part (e.g. WASTEWATER *is made of* CONTAMINANTS).

- *delimited-by*: this relation is used for physical objects, and marks the boundaries, dividing one object from another or marking the beginning or end of an entity. This is a domain-specific relation, mainly for geographic entities, such as the different layers of the atmosphere or the Earth. In wastewater treatment, however, it can be considered that the beginning of a WASTEWATER TREATMENT PLANT *is delimited by* the end of the SEWAGE SYSTEM.

- *located-at*: this relation is relevant when the location of a physical object is an essential characteristic for its description. For instance, a BREAKWATER is not a BREAKWATER if it is not located on the coast. This relation sometimes seems to converge with the *part-of* relation. In such cases, the *part-of* relation overrides the *located-at* relation. For example, a RIVER BED is *part-of* a RIVER instead of *located-at* a RIVER, because a RIVER cannot exist without its BED.

- *takes-place-in*: this relation describes the context of processes which have spatial and temporal dimensions. The distinction between this relation and *located-at* is based on the fact that processes are not as bounded in space as objects, and have a temporal dimension. For example, WASTEWATER TREATMENT *takes-place-in* the WASTEWATER TREATMENT PLANT; and THERMAL LOW *takes-place-in* SUMMER.

- *attribute-of*: this relation is only useful for concepts designated by specialized adjectives, such as ISOTROPIC, ALLUVIAL, ABYSSAL, etc., or nouns that designate the properties of other concepts, such as ALTITUDE, CAPACITY, COEFFICIENT, etc.

- *result-of*: this relation is relevant to either processes or entities that are derived from other processes. Even though processes and entities can be the result of another process, a process cannot be the result of an object. For example, EFFLUENT is the *result-of* WASTEWATER TREATMENT (process), but it cannot be regarded as the *result-of* a WASTEWATER TREATMENT PLANT (object).

- *affects*: this relation, along with *result-of*, are crucial conceptual relations in dynamic systems since both have a high combinatorial potential and can relate all kinds of concepts to changing environments. Actually, they are the only ones that can be shared by concepts belonging to all three semantic roles. They link processes or objects that cause a change in any other object or process without producing a concrete final result (e.g. SEWER PIPE *affects* WASTEWATER). Moreover, complex conceptual relations such as *affects* can generate a hierarchy of domain-specific relations such as *purifies*, *screens*, *compacts*, etc.

- *has-function*: this relation not only links objects or processes that are artificially created or carried out with a specific function, but also natural entities that, despite not being goal-directed, can be used for human profit. A natural concept with a function can be WATER (*has-function* IRRIGATION). As in the case of *affects*, *has-function* can also be associated with other domain-specific subordinate relations, such as *measures* for instruments (a PLUVIOMETER *measures* PRECIPITATION); *studies*, for sciences (WASTEWATER ENGINEERING *studies* WASTEWATER

TREATMENT); and *represents* for graphics, maps and charts (a HYDROGRAPH *represents* RATE OF WATER FLOW).

- *effected-by*: this relation is only used for instruments that carry out some process or create an entity. For example, SAND TRAPPING is *effected by* a SAND FILTER. This relation is especially meaningful in those domains where human interaction plays an essential role as is the case of environmental contexts.

The ontological implementation of conceptual relations can be seen in Figure 2, where the concept SEWER is described through three of them:

Figure 2. Concept SEWER in ontology

Those conceptual relations, specifically conceived for our Environmental TKB, can be enhanced by an additional degree of OWL semantic expressiveness provided by property characteristics, such as *transitivity*, *symmetry* and *inversion*.

In general terms, apart from the obvious transitive relation *is_a*, the *part_of* relation could also benefit from transitivity. Nevertheless, this can only be possible under certain conditions. Not all parts have the same implications with regards to their wholes and, in particular cases, transitivity may turn out to be misleading. For example, a RIVER HEAD is *part of* a RIVER and a RIVER is *part of* the HYDROSPHERE. However, a RIVER HEAD cannot be considered a *part of* the HYDROSPHERE, as it does not seem a very good example so as to how things relate in the real world.

Consequently, before considering transitivity, partonymy should be split up in several different relations which only apply for certain types of conceptual classes. At least, one relation for processes which are clearly divided in subprocesses (SCREENING<PRELIMINARY TREATMENT<WASTEWATER TREATMENT, so SCREENING<WASTEWATER TREATMENT) and another one for physical objects which are sharply bounded in space, as Figure 3 shows:

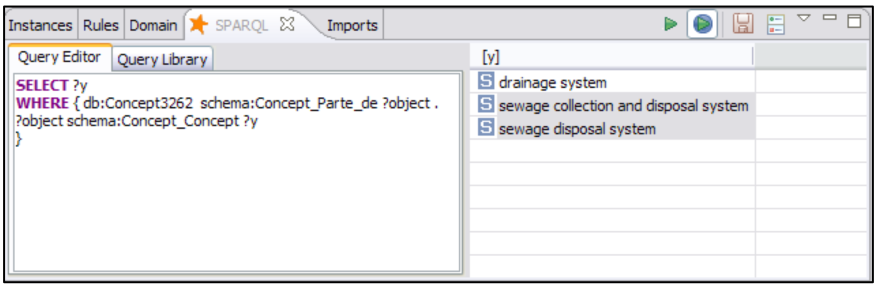


Figure 3. Inferred transitivity

On the left side, a SPARQL query is made in order to retrieve which concepts are *part of* Concept 3262, which refers to the concept SEWER. On the right side, DRAINAGE SYSTEM is retrieved as a direct *part-of* relation, whereas SEWAGE COLLECTION AND DISPOSAL SYSTEM and SEWAGE DISPOSAL SYSTEM are inferred through the Jena reasoner.

As for symmetry, the conceptual relation *delimited by* can be a good example. If the MANTLE is delimited by the EARTH CRUST, it can also be said that the EARTH CRUST is delimited by the MANTLE. On the other hand, all relations, except the symmetrical one, have its own inverse property (*result_of → causes; part_of → has_part; located_at → has_location; is_location_of; affects → is_affected_by, effected_by → has_instrument*, etc).

Finally, some examples of disjoint classes in our ontology are AGENT and PROCESS and PROCESS and PATIENT respectively, since they correspond to totally different notions. This is due to the fact that any physical object can be the result or the patient of any process and can even act as an agent in certain contexts (eg. WATER). On the other hand, a process can be the cause of a new action or the result of a previous one.

However, any specialized domain contains sub-domains in which conceptual dimensions become more or less salient depending on the activation of specific contexts. This means that concepts tend to be intermingled in different situations, and they are not always related to the same concepts or through the same relations, especially in such a wide domain as the environment, where entities are subject to constant transformation.

According to Barsalou [2005] a given concept produces many different situated conceptualizations, each tailored to different instances in different settings. Thus, context can be said to be a dynamic construct that triggers or restricts knowledge. Moreover, it has been claimed by Yeh and Barsalou [2006] that when situations are incorporated into a cognitive task, processing becomes more tractable than when situations are ignored, and so should be for knowledge acquisition processes. As a result, a more believable representational system should account for re-conceptualization according to the situated nature of concepts. This is done by dividing the environmental specialized field in different contextual domains. In this case, the WASTEWATER TREATMENT subdomain.

4. CONTEXTUAL RESTRICTIONS

Relational constraints have thus been defined according to concept types, multidimensionality and contextual factors. In this way users can browse the application from general to more specific context-based knowledge. This means that, at a broad level, concepts are described in a prototypical way and relational constraints only depend upon the main properties of each concept type (*entity*, *process*, etc) and the semantic role they possess. However, due to multidimensionality, one concept may have several roles (i.e. WATER can be an erosive AGENT or a PATIENT in water treatment processes) and thus behave in different ways. As a result, at a more specific level, concepts appear in more fine-grained representations, where context-dependent dimensions are added, restricted or highlighted.

The whole network of WASTEWATER TREATMENT covers several hierarchical nodes. In the first hierarchical level, the concept is obviously linked to only other context-specific concepts, but from the second hierarchical level, concepts should not show all dimensions, since there are many that come from a context-free database. Some of them are too general concepts and take part in very different events, and some others are still specialized concepts but are linked through different dimensions to very general ones, like SEWAGE COLLECTION AND DISPOSAL SYSTEM, which is directly linked to one of the upper levels in the hierarchy (SYSTEM).

But this mainly applies to non-hierarchical relations, which are the ones that reflect dynamism in the domain. For instance, SEDIMENT is a highly context-sensitive concept. Prototypically it can be first categorized as a *type of* DEPOSIT which is the *result of* a natural process like SEDIMENTATION. If the context of COASTAL DEFENSE arises, this adds a functional dimension as it is *used for* BEACH NOURISHMENT. Finally, in the context of WASTEWATER TREATMENT, SEDIMENT is the *result of* DECANTATION. Evidently, users would not acquire any meaningful knowledge if all dimensions of SEDIMENT were shown at the same time:

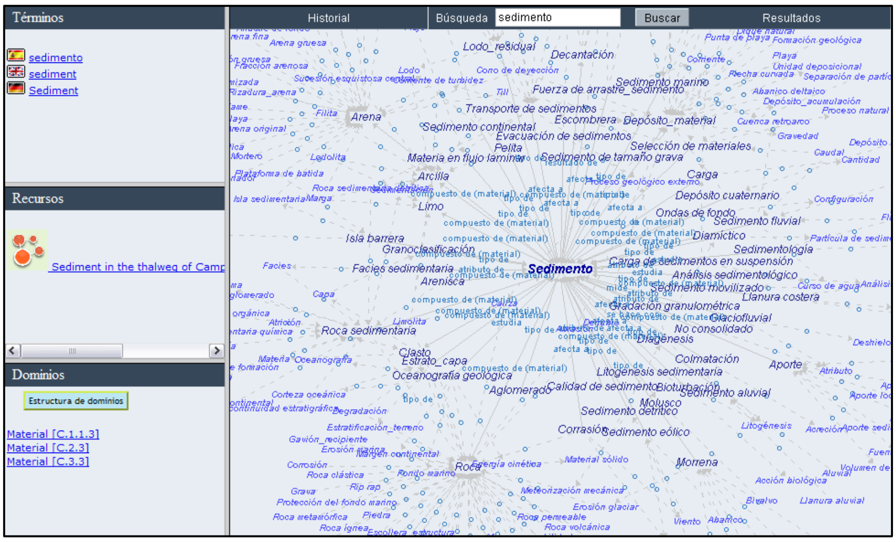


Figure 4. Network without constraints[1]

Thus, contextual re-conceptualization shows different relations associated to a particular concept depending on the activated context. Inferences in the ontology can then be optimized if they are performed according to different situations, which also tackle the problem of overinformation.

Consequently, when contextual restrictions are applied through the ontology, all concepts involved in WASTEWATER TREATMENT (no matter which hierarchical level they belong to) only show relevant dimensions:

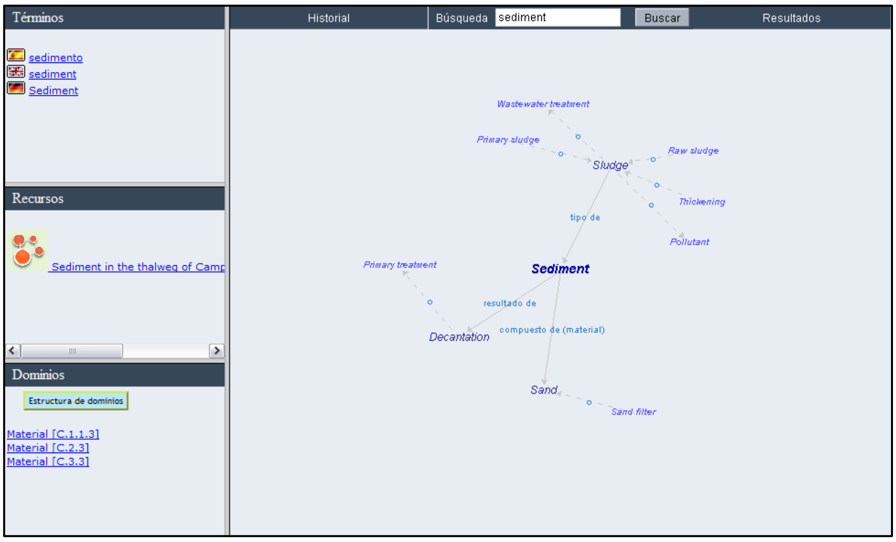


Figure 5. Contextual network[2]

[1] Spanish was used as a first reference language to represent conceptual information.
[2] However, we are currently working on the migration of the whole system into English.

5. CONCLUSIONS AND FUTURE WORK

We can conclude that legacy systems and ontologies can be integrated in a common framework by following a set of appropriate criteria. In addition, at this stage, linguistic semantics stored in the database can be enriched with OWL expressiveness. Nevertheless, the generalization of these procedures needs extensive further research, since they show a certain degree of application-dependency.

On the other hand, contextual restrictions have been proven to be a viable solution for managing overinformation and, at the same time, accounting for a cognitive-based representational system, which enables the TKB to meet knowledge acquisition requirements.

The next step in our project entails extending the procedure followed in the WASTEWATER TREATMENT context to other domains in the relational database. Efficiency is an issue that deserves special attention at this stage. Also, interconnection with outside sources of information is crucial. So, by taking advantage of the technologies used, methodologies of ontology mapping and matching can be applied.

ACKNOWLEDGEMENTS

This research has been partially supported by project FFI2008-06080-C03-01/FILO, from the Spanish Ministry of Science and Innovation and project P06-HUM-01489, from the Andalusian Regional Government. Pedro Magaña holds a FPU scholarship from the Spanish Ministry of Science and Innovation.

REFERENCES

Barrasa, J. “Modelo para la Definición Automática de Correspondencias Semánticas entre Ontologías y Modelos Relacionales” (PhD dissertation, Universidad Politécnica de Madrid), 2007.

Barsalou, L.W., Situated conceptualization, *Handbook of Categorization in Cognitive Science,* Cohen, H. & Lefebvre, C. (eds), St. Louis: Elsevier, 2005.

Bizer, C. & Seaborne, A., D2RQ-Treating Non-RDF Databases as Virtual RDF Graphs, *Proceedings of the 3rd International Semantic Web Conference (ISWC2004)*, 2004

Bowker, L, Multidimensional Classification of Concepts and Terms, *Handbook of Terminology Management*, 131-143, 1997

Faber , P. et al., Process-oriented terminology management in the domain of Coastal Engineering, *Terminology* 12, no. 2, 189-213, 2006

Faber, P., León Araúz, P. Prieto Velasco, J.A., Reimerink, A., Linking Images and Words: the description of specialized concepts (extended version), *International Journal of Lexicography,* 20:1, 2007.

Faber, P. et al. Semantic Relations, Dynamicity, and Terminological Knowledge Bases, *Proceedings of the XVII FIT Congress*, Shanghai, 2008.

Fillmore, C.J., Atkins, B.T.S., Towards a frame-based lexicon: the semantics of risk and its neighbours, *Frames, Fields and Contrasts*, Adrienne Lehrer y Eva Feder Kittay (ed.) Hillsdale, New Jersey: Lawrence Erlbaum Associates. 75-102, 1992.

Grinev, Sergei e Igor A. Klepalchenko, Terminological approach to knowledge representation, 147-151, 1999

Hřebíček, J., Legat, R., Nagyb, M., Current Trends in eEnvironment and its Role in eDemocracy, Conceptual Model of Single European Information Space for Environment, iEMSs Fourth Biennial Meeting: International Congress on Environmental Modelling and Software (iEMSs 2008). Barcelona, iEMSs, 1612-1619, 2008.

Meyer, I., Bowker, L. and Eck, K., COGNITERM: An experiment in building a knowledge-based term bank, *Proceedings of Euralex ’92*, 159–172, 1992

Yeh, W., & Barsalou, L.W., The situated nature of concepts, *American Journal of Psychology*, 119, 349-384, 2006

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

PEGASE
A Software Dedicated to Surface Water Quality Assessment and to European Database Reporting

Ir. J-F. Deliège, Ir. E. Everbecq, Ir. P. Magermans, A. Grard, T. Bourouag, C. Blockx
Aquapôle, University of Liège, Belgium (aquapole@ulg.ac.be)

Abstract : The Aquapôle of the University of Liège has been involved in environment modelling for more than 20 years. Its R&D unit focused on the domain of surface water, targeted to water quality. One of its environmental models, called PEGASE (Planification Et Gestion de l'ASsainissement des Eaux), French acronym for "Planning and management of water purification" is devoted to the characterization of the environmental state of surface water, at the scale of a whole watershed, basin and district. The software has been used operationally in several countries like France and Benelux, but several calculations have also been performed at international level for the international commissions in charge of the Scheldt and the Meuse.

PEGASE model requires as input, at least data like digital terrain models, water flow, water level and other hydrodynamic measurements in some geo-localized points, diffuse loads properties, ecological model data (to characterize bacteria, phytoplankton, zooplankton, macrophytes, shells, etc…), chemical discharges (from industry and cities) and treatment plant effect. The output of the software consists in tables, graphs and maps, showing the state of various calculated parameters of quality (DO, BOD, COD, N, P, Chla,… + indices) either at a specified time, along the river, or the temporal evolution at a given point. It provides printed maps or animations of the evolution of the results on the basin.

The software is able to simulate non stationary scenarios from a local scale, up to a global scale i.e. a few km² to hundreds of thousands km² and handles the entire river tree (hundreds to thousands of water bodies). Some results of the model are already used to populate the SoE-Wise databases. On another hand, the model is a tool to validate and check consistency of SoE-Wise data at district (and thus international) level. The flexibility of the model enables to choose the time (daily to yearly periods) and spatial (watershed, water body, district, region, country, …) scales of the expected results.

PEGASE is also a full-scale operational tool for WFD implementation and WISE databases. This means that our partners handle a software (also known as the model) directly usable – and of direct help – in the implementation of the WFD and other water related directives. The model allows the member states to compare efficiently the collected international WISE data and populate database with relevant model results. We propose not only to read the data directly from the database to avoid manipulation mistakes, but also we propose to generate output data in a format directly exploitable by WISE databases administration.

Keywords: environmental data; international database; environmental modeling; water quality assessment

1 INTRODUCTION

The Aquapôle is a research centre, in charge of developing expertise and knowledge in all water related domains. A number of services within the University of Liège, and other Universities are members of the Aquapôle. The Aquapôle is located on the campus Sart-Tilman of the University of Liège (Belgium). It is an interface between the university and external public and private or industrial world at local, national and international level.

Its own R&D department currently focuses on several topics:

- the integrated water analyse and management in a perspective of sustainable development by
    - the improvement of the existing management tools, by taking into account ecological, economic and social stakes;
    - the development of predictive models to help management (quality river model PEGASE, and integrated models of the hydrologic cycle such as MOHISE, MOHICAN, SALMON, MOIRA);
    - supporting the public and private operators in the implementation of the European Directives (water 2000/60/CE, nitrates, ...);
- the treatment and management of sewage effluents;
- analytic measurements of water quality;
- the cooperation with Southern countries;
- the knowledge of aquatic ecosystems and impacts related to the pollution.

The main partners for this latter are the public authorities of various European countries, more particularly, the authorities in charge of surface water management.

For these partners, the Aquapôle has developed a software called PEGASE (Planification Et Gestion de l'ASsainissement des Eaux) devoted to the characterisation of the environmental state of surface water, at various scales from local watershed to the whole district. The software has been implemented and is operationally used in several countries like France, Belgium and Luxemburg, but also several calculations have been performed for transnational basins for the international commissions in charge of the Scheldt and the Meuse district.

The model, originally developed to help water decision-makers to meet their obligations, is also able to help them to collect and consolidate the information necessary for international reporting (e.g., WISE)

2 MODELLING OF SURFACE WATER QUALITY AND SIMULATION

2.1 Context

The management of the aquatic resources is a particularly complex problematics, which involves many interrelated aspects. Decision making depends on competitive objectives and thus require a multi-compartment, global-scale and synoptic approach. In this domain, an approach based on mathematical models is relevant and highly recommendable. E.g., the PEGASE model allows assessing the effects of some laying out policies, and provides the water decision-makers with a water system analysis tool for seeking alternative or more efficient solutions.

The purpose of PEGASE model is to help make decisions – as regards to surface water management – by the calculation of water quality according to pollutant contributions and releases, and of hydrological conditions.

The PEGASE model was built with the following motivations:

- to better understand the aquatic environment mechanisms (hence the construction of a deterministic and physically based model which requires very few calibration to be applied to new basins);
- to structure knowledge (including the "data input");
- to quantify the physically based "pressure – impact" relations and to help the authorities in charge of the surface water management in their process of decision making;
- to extend the "rivers" models to explicitly take into account the influence of their watershed.

PEGASE is an integrated model "basin/river" which makes possible to calculate in a deterministic way the water quality of the rivers according to pollutant contributions and releases, for various stationary hydrological situations. Obviously, the model can also operate in non stationary mode over several years. It can also provide an estimated calculation of the water quality improvements resulting from purifying actions or release reductions. The model includes a hydrological and hydrodynamic sub-model, a thermal sub-model, a sub-model dedicated to releases, and a sub-model dedicated to calculate the water quality and to explicitly describe the aquatic ecosystem mechanisms.

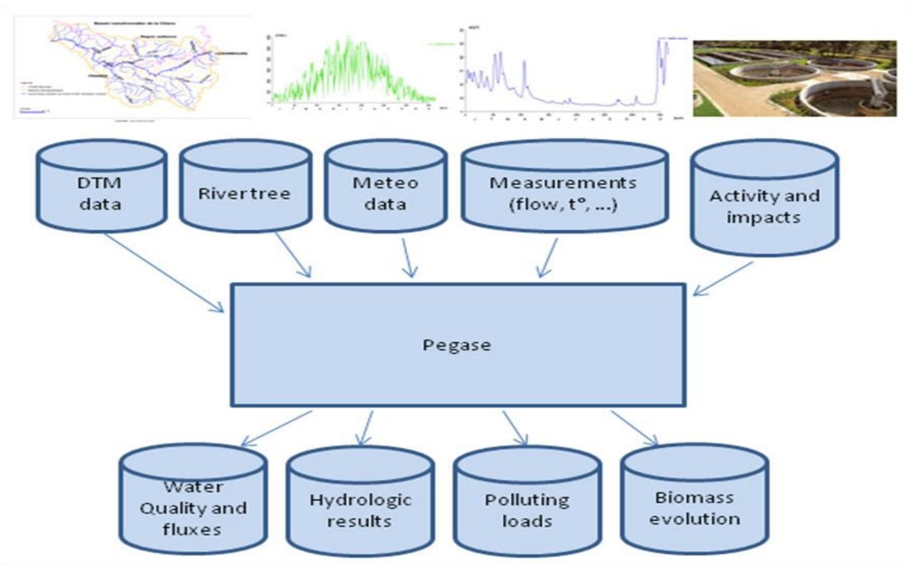


Figure 1 — PEGASE data flow

PEGASE was initially developed for the Walloon Region (at the end of the eighties). It then was adapted, improved and applied to other river basins. A special attention is paid by the developers to

(1) the upgrading possibilities (continuous improvement),
(2) the programs' modularity, and
(3) the needs and feedbacks of the users (existence of a user community).

The relevant utility of the PEGASE model was further increased in 2000 with the setting up of the European Directive 2000/60/CE establishing a framework for Community action in the field of water policy. Currently, PEGASE is used in the Walloon Region, but also in France (four water agencies), in the Flemish Region, in the Luxembourg and for several international basins (the Meuse, the Scheldt, the Mosel-Saar). Some applications were also tested on the upper Wisla river (Poland), the Nicolet River (Canada) or the river basin of Itajai (Brazil).

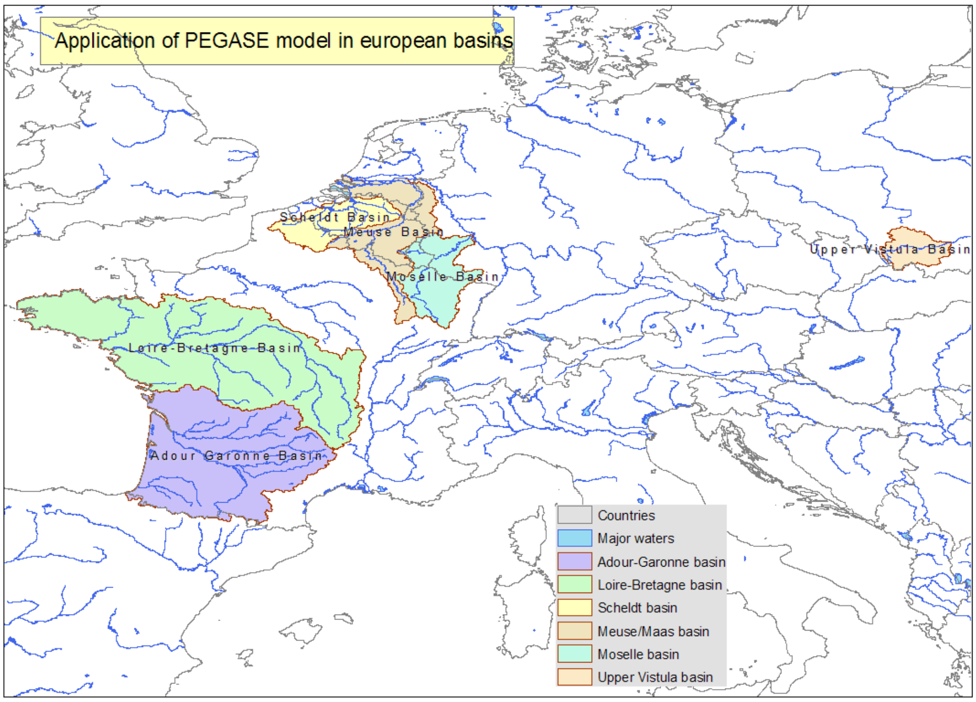


Figure 2 — PEGASE model applications in Europe

2.2 Input Data

PEGASE model is fed by several kinds of input :

- **Geographical data:** the PEGASE model can process several hundreds of rivers simultaneously and the associated surface of the river basins that can reach several tens of thousands km² (see figure below: the Adour-Garonne river basin). PEGASE also allows carrying out refined simulations on a subset of the entire river network (for example, on a single river having a river basin of a few tens km²).

  Among others, some input data are:
  - digitalized water courses;
  - digital terrain models;
  - land cover;
  - livestock;
  - administrative reference frame.
- **Hydro-meteorological data:** PEGASE uses measurements of daily water flows, daily temperatures, and insolation.
- **Data related to human activities and releases** (industrial, urban, water treatment plants, ...)**:** PEGASE describes in a structured way the urban releases, the industrial releases, the role of the purification plants, the releases due to livestock farming activities, and the diffuse contributions by the soils;
- **Other data** such as water level, hydrodynamics, anthropic singularities, ...

The input data can be processed and visualised using Geographical Information System (GIS).

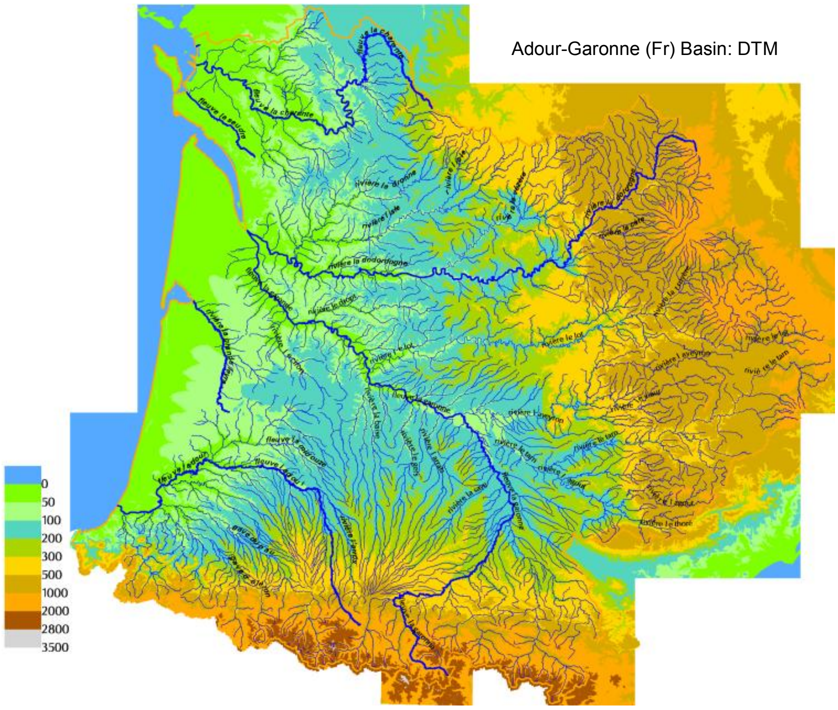


Figure 3 —Example of a river basin of several tens of thousands km²: the Adour-Garonne basin

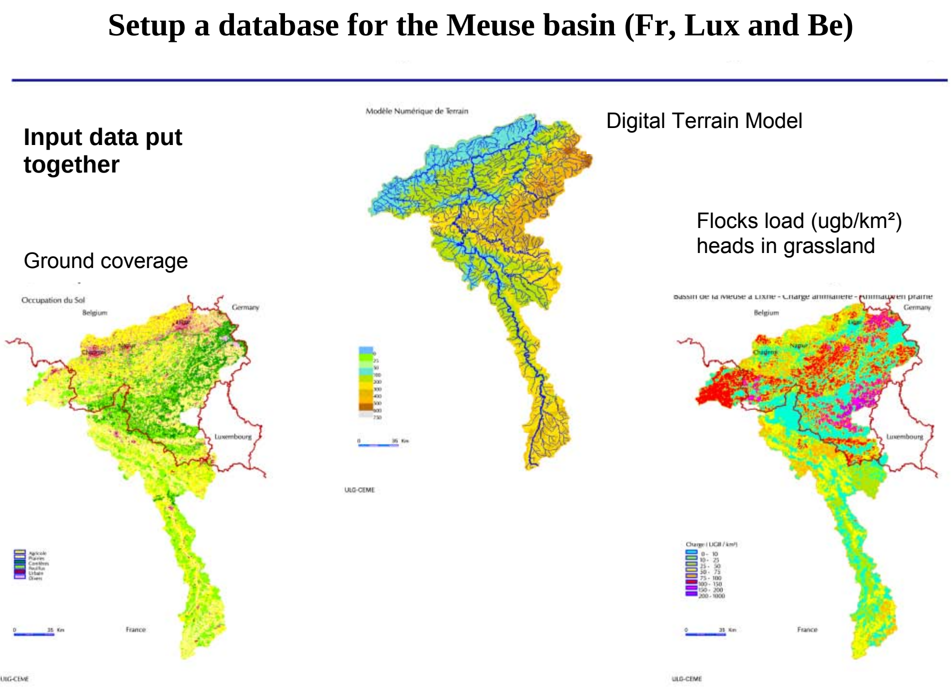


Figure 4 — Example of pre-processing on the French, Belgian and Luxemburger part of the Meuse basin (amongst 5 countries and 9 Competent Authorities for the whole basin)

2.3 Simulations and Results

The main processes and functionalities integrated into PEGASE model are:

- the explicit calculation of the primary production (phytoplanktonic biomass) using a multi-species sub-model;
- the explicit calculation of the self-purification mechanisms in the rivers and the evolution of eutrophication (calculation of the bacterial biomasses and their activity);
- the calculation of the micro-pollutants (heavy metals,…) ;
- the calculation of physicochemical quality indices (for example French SEQ-EAU);
- the estimate of biological quality indices;
- a module of cost/efficiency analysis.

PEGASE can be used in two modes:

- the “stationary” mode, in which the hydro-meteorological conditions (flows, temperatures, insolations) are considered as representative. This mode permits carrying out “simple” simulations which are easy to analyse and compare with other simulations;
- the “non-stationary” mode, in which the hydro-meteorological conditions vary hourly and thus permits calculating annual evolutions of the concentrations on the whole modelled hydrographical network.

Then, the PEGASE model allows comparison of scenarios in order to determine the optimal policies.

The output of the software consists in tables, graphs and maps, showing the state of various parameters of quality (DO, BOD, COD, N, P, Chla,… + indices) either at a specified time, along the river, or the annual evolution in a given point. It provides maps, or animations of the evolution of the results and calculations on the basin. The output results can be visualised using Geographical Information system (GIS).

Validating simulations are first realised for each river basin.

Figure 5 shows an example of validation: the annual evolution of simulated and measured nitrogen variables.

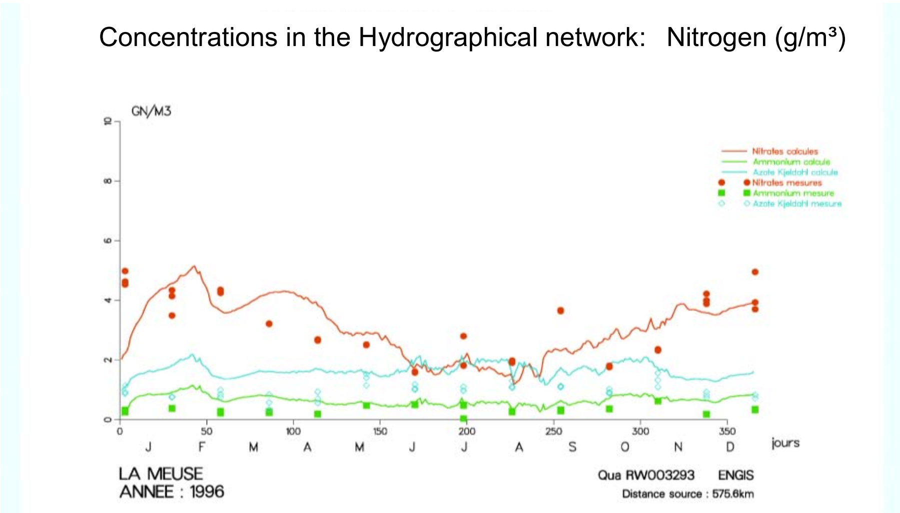


Figure 5 — Example of comparison on discrete measure points for calibration/validation

3 PEGASE AND THE INTEGRATING ENVIRONMENTAL KNOWLEDGE IN EUROPE

3.1 European Databases

The European Environment Agency (EEA) is a European public body in charge of providing objectives, reliable and comparable information on the environment. Since 1990, the EEA develops the European environment information and observation network (EIONET). The EEA is also responsible for the setting of the Water Information system for Europe (WISE). This information system is expected to provide the necessary data and information to facilitate their integration with the water related directives (WFD, UWWD, BW, ...). WISE, still under development, is already used for some data report and is partly available to the public on the EEA website.

Member states come up against national datasets that do not fit to WISE datasets. For this reason, not only are we proposing to read the data directly from the national databases to avoid manipulation mistakes, but also we propose to generate output data in a format directly exploitable by the WISE databases. WISE aims to simplify reporting of EU water-related information, and enhance use of that information, but new tools are essential at end user level to reach that goal.

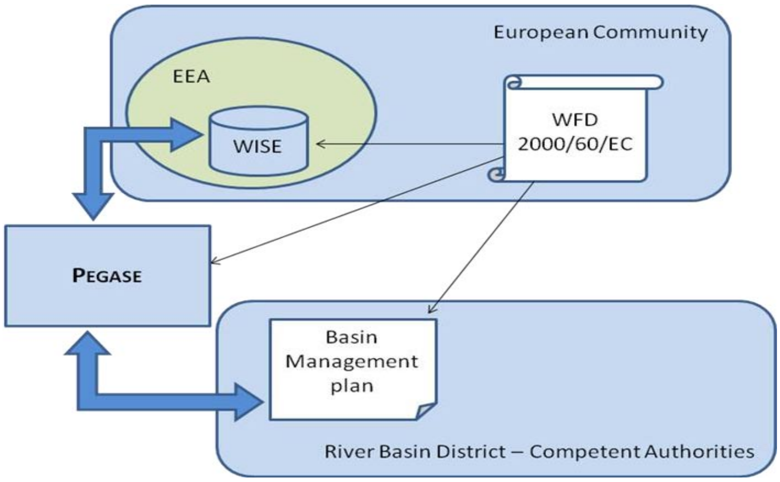


Figure 6 — PEGASE within the European Environment

3.2 Evolution of PEGASE with respect to EIONET

The wish for the coming projects is to bring PEGASE to a full-scale operational level for WFD implementation and WISE database usage. This means provide to our partners a software program directly usable, and of direct help in the implementation of the WFD. The model will also allow the member states to use efficiently the collected international WISE data and populate database with relevant model results.

A LIFE+ project in that way has been submitted last November under the reference LIFE08 ENV/B/000041 ("Add WFD/WISE compliance to a software suite dedicated to Surface Water Quality Assessment"). If this project were accepted, we would demonstrate the benefits provided by the extension of PEGASE features and outputs thanks to specific use cases e.g. Scheldt and Meuse: Two international rivers, which are currently modelled using PEGASE might serve to demonstrate the added value, by directly comparing the output of PEGASE prior and after the modifications. The results would be submitted to the international commission in charge of the respective basin. We also intend to demonstrate the transposability to other regions, and the usability of the results by other international commissions (or sub-commissions) by studying another use case situated in eastern Europe, where no calculation has been performed with PEGASE earlier (candidate basin still to be identified).

4 CONCLUSION

PEGASE proves to be an efficient tool for helping in surface water management up to the international district level. Already used by several basin management competent authorities, the model structures the river ecosystem knowledge and represents:

- A new way to handle the assessment of the quality of the rivers, by non-stationary, accurate and physically based calculation on the whole basin, and comparison on discrete measurements points for calibration/validation. These results lead to a much more precise knowledge of the rivers network quality with a detailed discretisation of the network (e.g. 1 node every 200 m or less) for annual simulation periods.
- A new way to assess the impact of measures taken to enhance the quality of targeted water bodies, by performing simulations of scenarios on the basin, before physically building anything, and thus assess the cost/efficiency ratio of each solution.
- A new operational way to assess the impact of climatic changes on the future river quality, allowing planning of measures several years before they are needed, allowing optimised lifecycle and cost in the realisation of these measures.
- A new way to ensure the consistency of the data at international level.
- A new way to extrapolate discrete measurements (in time and space) to each water body by a sophisticated physically based calculation.
- A computer tool adapted to the new requirements imposed by the issue of the WFD and WISE, providing directly usable results to the decision makers.

APPENDIX — TABLE OF ACRONYMS

Acronym	*Meaning*
BOD	Biological Oxygen Demand
BW	Bathing Water
Chla	Chlorophyll A
COD	Chemical Oxygen Demand
DO	Dissolved Oxygen
EEA	European Environment Agency
EIONET	Environment Information and Observation Network
EU	European Union
GIS	Geographical Information System
PEGASE	Planification Et Gestion de l'ASsainissement des Eaux
SoE	State of Environment
UWWD	Urban Waste Water Directive
WFD	Water Frame Directive
WISE	Water Information System for Europe

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Advanced corporate sustainability reporting – XBRL taxonomy for sustainability reports based on the G3-guidelines of the Global Reporting Initiative

R. Isenmann[a] **and J. Marx Gómez**[b]

[a] *Fraunhofer Institute for Systems and Innovation Research (ISI), Breslauer Straße 48, 76139 Karlsruhe, Germany (ralf.isenmann@isi.fraunhofer.de)*

[b] *Carl-von-Ossietzky-Universität Oldenburg, Business Informatics I, Ammerländer Heerstr. 114-118, 26129 Oldenburg, Germany (jorge.marx.gomez@uni-oldenburg.de)*

Abstract: Sustainability reporting describes a development path towards a concept of balanced corporate reporting, usually communicating the three pillars of environmental, social, and economic performance and its mutual interrelations, what in business terms is often called the triple bottom line approach. While early sustainability reports merely have been available on print media, today most are accessible on the WWW as HTML files or as PDFs. Such a layout oriented data format however, does not seem to be sufficient any longer, especially in terms of content-syndication, harmonization, efficiencies, future ICT needs, stakeholders' expectations, and data exchange inside and outside companies, and – no less important – in the light of approaching a European system for collecting, analyzing, and reporting environmental information (SEIS). In order to meet these challenges, a standardized reference architecture for sustainability reports built on the eXtensible Businesses Reporting Language (XBRL) is proposed, and its underlying methodology is described. The entire development is based on a research initiative embedded in the German Environmental Informatics Community. The proposed document structure particularly meets the requirements of the Global Reporting Initiative's (GRI) sustainability reporting guidelines (G3). While developing a unifying document structure for sustainability reports, we made use of on existing XBRL Financial Reporting Taxonomies Architecture (FRTA) which is common for business financial reporting. Such a document structure is key for advanced reporting systems, especially current ICT applications like (web) content management systems. Using the reference architecture offers an impressive array of benefits, e.g. it helps: to facilitate data exchange between reporters and report users, to improve a company's information management, supports its reporting workflow, allocates its resources efficiently, exactly meets different requirements proposed by emerging guidelines, and refines communication with its target groups in a meaningful way. As a larger goal, the whole methodology which has been developed in document and software engineering could be transferred to Chemical Safety Reporting and its currently emerging requirements, based on the European Commission Regulation concerning registration, evaluation, authorization and restriction of chemicals (REACH). For this reason, if no other, all stakeholders involved in Chemical Safety Reporting should benefit from prior experiences in the closely related field of sustainability reporting.

Keywords: Chemical Safety Reporting; REACH; Sustainability Reporting; eXtensible Business Reporting Language (XBRL); Global Reporting Initiative (GRI).

1. INTRODUCTION

Corporate sustainability reporting has its roots in environmental respectively in non-financial reporting (IISD et al., 1992; DTTI et al., 1993; UNEP and SustainAbility, 1994). It follows a development path towards a concept of balanced reporting, usually communicating the three pillars of environmental, social, and economic performance and its mutual interrelations, in business terms often called the triple bottom line approach (Elkington, 1997). Sometimes, this approach is put in popular terms like “making values count” (ACCA, 1998), or “linking values with value” (KPMG, 2000), or described as “creating value and optimizing prosperity according to the Triple P bottom line” (SER, 2001; DCCA, 2006). The latter is understood as combining shareholder value, eco-efficiency, and corporate citizenship, or being part of corporate social responsibility (CSR Europe, 2000).

In the 10 years since sustainability reporting first became a topic of broader interest in academia, business, and government, it has rapidly grown to a field of research with increasing relevance for companies (Kolk, 2004) and capital markets (Australian Government, 2003; Hesse, 2007), even in the eyes of investors (Australian Government, 2003; BSR, 2008). At present, sustainability reporting seems to become part of companies’ daily affairs, even entering (to a certain extent) the business mainstream. Hence, for a growing number, not just for some pioneering companies, the question is now how to report on sustainability issues, and no longer whether to report at all (Marshall and Brown, 2003). In parallel, a solid and powerful institutional infrastructure for corporate sustainability has been built, with various initiatives and organizations in Europe (Waddock, 2008).

Regardless of nationality or other differences in country results, this is not only true for leading edge companies in corporate sustainability and few sector leaders, but also for global players and multinationals (KPMG, 2008), stock-quoted and publicly traded companies (Raar, 2002), as well as for a number of medium-sized (Clausen et al., 2001) or small companies (EC, 2002). This trend is evidently a worldwide phenomenon (KPMG, 2008), with Europe and North America coming first, followed by the Asia-Pacific region, and even spreading to Africa (Visser, 2002).

Within several industrial sectors, there is further empirical evidence that environmental and sustainability reporting today has become of competitive relevance (Fichter, 1998) and strategic importance (Larsen 2000), with an impact on brand value (Interbrand, 2008). Today, “greenwashing” (Futerra, 2008), i.e. merely provision of “green glossy brochures” (UNEP and SustainAbility, 1994), does not seem to be sufficient any longer; a substantial amount of information is required. Further, sustainability reporting is only successful if the underlying management systems are appropriate and the associated processes are effective and operational. For example, goals have to be set, responsibilities have to be assigned to reach the goals, and outcomes must be assessed and used as the basis for forthcoming efforts.

Following Mesterharm (2001), comprehensive environmental or sustainability reports are regarded as the primary and leading vehicles and thus the pivotal instruments of such voluntary communication (Brophy and Starkey, 1996) because of its unique claim to credibility and reliability external stakeholders ascribe to it, containing quantitative and qualitative data. These reports are usually addressing a wide range of target groups, are often produced as single documents and issued for a certain period of time. Companies use such reports for disclosing activities and integrated performance, often including the following topics: top management statement, management policy and system as well as input-output-inventory of impacts of production processes and products in terms of sustainability.

While the field is still evolving, as sustainability reporting matures and practice develops into a more sophisticated stage, companies have to realize that the “honeymoon period” (DTTI et al., 1993) in which comprehensive non-financial reports received media and public attention just for the fact that they publish reports at all rather than for what was disclosed is over. Nowadays, advanced reporting approaches with substantial information are required. However, further to the relevance of contents, issues of communication style and data quality also become of greater importance (Beattie and Pratt, 2003; Hund et al., 2004; ACCA, 2004), in particular:

- interactivity (Teo et al., 2003; Isenmann and Kim, 2006),
- target group tailoring (Jensen and Xiao, 2001; Isenmann and Marx Gómez, 2004),
- and stakeholder dialogue (WBCSD, 2002; Unerman and Bennett, 2004).

Due to cross media availability and other innovative opportunities offered by the internet and its associated technologies and services, companies are entering a new transitional stage of online reporting (SustainAbility and UNEP, 1999; Clarke, 2001; Wheeler and Elkington, 2001; Isenmann et al., 2007).

For example, in “The 2001 Benchmark Survey of the State of Global Environmental and Social Reporting” carried out by the CSR network (Line et al., 2002), internet-based reporting and a more balanced reporting approach are seen as the top reporting priorities. Just a short time later, many sustainability communication vehicles and reporting instruments are already available on the WWW, or – at least – benefit from internet support (ACCA, 2001; Shepherd et al., 2001; Isenmann and Lenz, 2002; Scott and Jackson, 2002; Rikhardsson et al., 2002; Andrew, 2003; Lodhia, 2004; Isenmann, 2004): Reports, brochures, leaflets, newsletters, press releases, slides, presentations, audio sequences, video clips etc. are accessible via download and/or online, or can be “pulled” or automatically disseminated via email or other current “push” technologies (Isenmann and Lenz, 2001). Despite progression companies have made in recent years however, reports are more or less available in a layout oriented data format like HTML and PDF. Such a standalone disclosure practice does not seem to be sufficient any longer, especially in terms of content-syndication, harmonization, efficiencies, future ICT needs, stakeholders’ expectations, and data exchange inside and outside companies, and – no less important – in the light of approaching a European system for collecting, analyzing, and reporting environmental information (SEIS).

This paper describes the development of a standardized reference architecture for sustainability reports using the eXtensible Businesses Reporting Language (XBRL) and its underlying methodology. The development is based on a research initiative embedded in the German Environmental Informatics community. The proposed unifying document structure particularly meets the requirements of the Global Reporting Initiative’s (GRI) sustainability reporting guidelines (G3) released in October 2006. While developing the document structure for sustainability reports, we made use of on existing XBRL Financial Reporting Taxonomies Architecture (FRTA) which is common for business financial reporting. Such a document structure is key for advanced reporting systems, particularly for current ICT applications like (web) content management systems. Using the reference architecture offers an impressive array of benefits in corporate reporting, e.g. it helps: to facilitate data exchange between reporters and report users, to improve a company’s information management, supports its reporting workflow, allocates its resources efficiently, exactly meets different requirements proposed by emerging guidelines, and refines communication with its target groups in a meaningful way.

As the overall aim, the experience gained in sustainability could be transferred to Chemical Safety Reporting. Particularly methodological insights we have made in document engineering are of particular relevance for all stakeholders involved in the up and coming new regulation called: REACH (EC, 2008). These insights may provide an excellent source for meeting the European Commission Regulation concerning registration, evaluation, authorization and restriction of chemicals.

2. FRAMEWORK FOR ADVANCED SUSTAINABILITY REPORTING

A framework for advanced sustainability reporting could be illustrated along four elements (Isenmann and Marx Gómez, 2004, 2008; Isenmann, 2005):

- stakeholder analysis (section 2.1),
- analysis of stakeholder information requirements (section 2.2),
- XML-based document engineering (section 2.3) and
- ICT-architecture (section 2.4).

The framework may serve as a guideline on how to exploit latest media-specific capabilities that the internet and its associated technologies and services provide, surely leading to improvements in the area of sustainability reporting, but also applicable to Chemical Safety Reporting.

2.1 Stakeholder analysis

The starting point of any advanced sustainability reporting system is a stakeholder analysis identifying the primary users and typically asking: Who are the relevant stakeholders (including the critical ones), which key target groups inside and outside the company require information via environmental reporting (Figge and Schaltegger, 2000)? Generally, there are two ways of identification, either a deductive approach or an inductive approach combined with a deductive approach.

According to the deductive approach, initially all stakeholders could be considered relevant or called a target group who are involved in or affected by a company's environmental impacts and activities. Perhaps, as certain stakeholders claim some exclusive information rights, they may be seen as specific users. For example, this is true for companies' top managers who hold ultimate liability, for local authorities who have a specific right to know and also for banks and insurers who require confidential information. Regardless of certain information rights, it could be fruitful anyway addressing these groups as users, too.

Despite its proven usefulness, the deductive approach should be combined with an inductive one for this task. Stakeholder analysis represents a company-specific task influenced by certain circumstances, e.g. size, industry, products, processes, location, environmental impacts, stakeholder relations, communication strategy, environmental management, and strategic goals. Hence, an empirical analysis could validate the number of relevant stakeholders that are found through the deductive approach. Lenz (2003) provides a comprehensive stakeholder analysis. He reviewed a multitude of empirical studies that identified key target groups and primary users' needs in the field. Resulting from his in-depth survey he identified 12 key target groups, arranged in a stakeholder map with four groups:

- Financial community, including investors, insurance agents and financial analysts;
- business partners, including employees, customers and suppliers;
- diffuse groups, including media representatives, neighbors and consultants;
- normative groups, including local authorities, respective legislators, pressure groups and standard setting institutions.

To some extent, the users within a certain target group can be indicated through relatively homogeneous information needs.

2.2 Analysis of stakeholder information requirements

Following stakeholder analysis and identification of primary users, a reporting organization should study information needs and other preferences expected to be met in report form and content. Such an analysis of stakeholder information requirements are meant to determine relevant contents that target groups expect and the preferences they require regarding form, layout, design, media and distribution channel. There is consensus that meeting users' needs is needed for successful elaborated environmental and sustainability reporting (Isenmann and Kim, 2006). In contrast to its wide acceptance in frameworks and guidelines, however, current practice shows another picture, with significant room for improvements, even for the best reporters. At present, little work has been done to conceptualize users' information needs, especially concerning distribution channels, presentation styles and media preferences (Pleon, 2005). Hence, van Dalen (1997) mentions a lack of more profound insights on users' information needs and preferences. Answering this need, Lenz (2003) provides a considerable analysis on stakeholder information requirements.

- For example, employees are interested in environmental and social performance of their employers and companies. They want to be informed about targets and activities related to the environmental management system. Further, they want to understand how companies are seen by local community groups. Employees wish to see their company as a going concern, recognizing that environmental performance might have some influence on this.
- In supply chains and other manufacturing networks, suppliers exchange information with participating business partners (Lippman, 2001). Establishing partnerships implies extensive environmental communication along the whole supply chain or network. These groups need environmental information regarding resource efficiency, regulatory compliance, new product and service opportunities, especially in terms of extended product stewardship, and other environmental liabilities.
- Investors, including institutional and private shareholders, financial analysts and investment consultants, are increasingly interested in environmental issues and their financial interrelations since they notice that sustainability reports make good business and environmental sense (Hesse, 2007). A number of investors expect that environmental performance influences financial performance and shareholder value. For example, in November 2000 a group of 39 financial investors, managing combined assets excess of $140 billion, sent a letter to CEOs of the 500 largest U.S. companies urging them to provide sustainability reports (SocialFunds, 2000).

Together, the analysis of stakeholder information requirements clearly demonstrates that employees, customers, suppliers, local authorities, legislators, neighbors, consultants, financial analysts, investors, insurance agents, media representatives and members of rating and ranking organizations have heterogeneous information needs. These different needs cannot be fully satisfied or easily met just by "reporting as usual" through orthodox practice, via one universal document (on print media), mostly produced as a "one size fits all" report. Users expect more and more customized reporting instruments (Isenmann et al., 2003). It is necessary to determine what target groups want, to identify their needs and preferences.

The results of the two analyses discussed above lend themselves to the creation of specific user profiles. For each of the core target groups, a profile of their information needs will be established that comprises content requirements, preferences as to the reporting form and secondary requirements such as the distribution channel.

2.3 XML-based document engineering

The results of stakeholder analysis and deeper insights of stakeholder information requirements are used for XML-based document engineering, indicating the ICT-laden area where contents, structures, procedures and design of reporting instruments and other communication vehicles are defined. This leads to the questions: What should an advanced sustainability report prepared as an XML document look like? What contents should be included? Who should be addressed? On what devices should the report be available? Which standards or guidelines need to be adhered to? Here, certain aspects of report structure, contents and layout are explicitly considered.

The core of XML-based document engineering is to develop standardized document structures for a certain type of documents (i.e. sustainability reports). This is usually done in the form of an XML schema, or more advanced, in the form of an XBRL taxonomy (eXtensible Business Reference Language): A schema defines the semantics and overall pool of contents in a basic structure for a certain group of documents, in this case, for sustainability reports. Of this pool of structured contents, customized reports exactly meeting the requirements of certain user, user groups or guidelines respectively can be prepared in an automated fashion by machine processing. In terms of document engineering, a schema consists of several elements representing the contents and their corresponding attributes, specifying the semantics and indicating the elements. Consequently, a schema determines what elements can be used within a XML document. Further, a schema describes how elements can be arranged, and which attributes certain elements may carry.

Embedded in a research initiative in the German Environmental Informatics community, the existing XML-schema for sustainability reports is blended in a standardized reference architecture for sustainability reports built on XBRL (Isenmann et al., 2008; Arndt et al., 2006). This proposal is developed on three basics:

- In terms of materiality and contents, the XBRL-taxonomy meets the requirements of the sustainability reporting guidelines (G3) of the Global Reporting Initiative (GRI, 2006).
- In terms of design and configuration, an existing XBRL Financial Reporting Taxonomies Architecture (FRTA, e.g. Hamscher, 2005) is applied. As the FRTA is common for business financial reporting, it helps that the XBRL taxonomy for sustainability reports may be used in the well-established field business financial reporting.
- In terms of underlying methodology, a generic process-model for the development of schemas was used.

2.3.1 G3-Guideline of the Global Reporting Initiative

The GRI's G3 specifies basic contents of a sustainability report which are of relevance to reporters and report users (GRI, 2006):

- Strategy and profile items set the overall context for understanding organizational performance such as strategy, profile, and governance;
- management approach items cover how an organization addresses a given set of issues to provide context for understanding the performance in a certain area;
- performance indicators provide comparable information along the economic, environmental, social, and integrative performance.

2.3.2 Financial Reporting Taxonomies Architecture

The Financial Reporting Taxonomies Architecture (FRTA) provides a reference architecture for business financial reporting. It is characterized through a number of so-called taxonomy schemas and linkbases. A taxonomy schema is an XML schema, usually a standardized schema (XSD) and contains XBRL concepts. A concept is a definition on how to report about the activities or nature of a business entity. Taxonomies contain XBRL concepts represented by XML schema element definitions.

Using the design principles of the FRTA delivers specifications for a proper development and overall structure of a XBRL taxonomy for sustainability reports. These specifications are crucial for data exchange and information management as they ensure consistency between different XBRL taxonomies at different levels, e.g. in terms of correct representation, modularity, and evolution.

2.3.3 XBRL taxonomy for sustainability reports

Following the FRTA, a XBRL taxonomy for sustainability reports consists of five elements (fig. 1):

- The XML schema is the core and represents the pivotal document of the taxonomy for sustainability reports. The schema contains all GRI disclosure items which are represented as XML elements and specified through certain data types.
- The labelLinkBase links all XML schema elements to certain literal meaning. For example, labels in English and German may be assigned to all XML schema elements if the sustainability report should be available in two languages.
- The definitionLinkBase defines a hierarchical order within the labelLinkBase how to structure the overall contents and where to fix certain XML schema elements in the sustainability report.

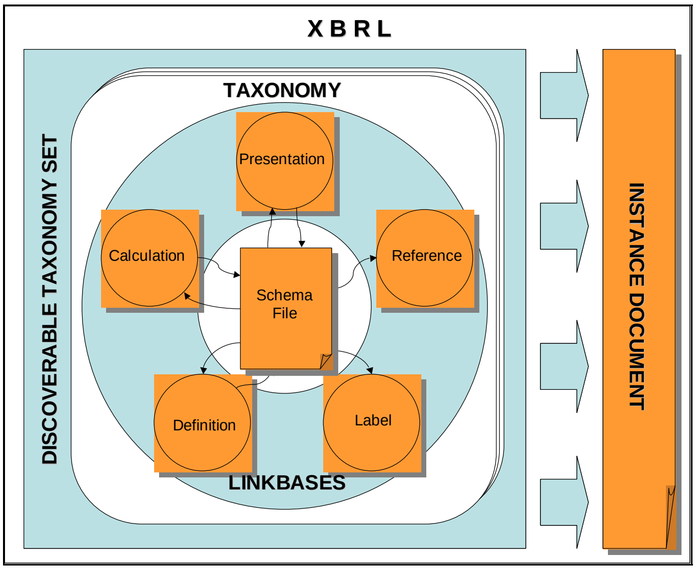


Figure 1: XBRL framework

- The presentationLinkBase makes it possible to arrange elements in any hierachical order without changing the overall structure defined in the definitionLinkBase.
- The referenceLinkBase provides the opportunity to link XML schema elements with further information like regulations, guidelines, and comments, and hence to facilitate the overall understanding.

According to the XBRL framework, there is a sixth conceptual element, i.e. calculationLinkBase. Up to now, however, the GRI-G3 has not provided a method for the calculation of performance indicators. The calculationLinkBase defines rules on how to calculate performance indicators and other data.

2.3.4 Process-model used as underlying methodology

The development of a schema for sustainability reports is a sophisticated work as various different requirements and heterogeneous needs have to be taken into account. The process-model used here was initially proposed by Schraml (1997) and then refined and elaborated by Lenz (2003) and Brosowski et al. (2004). It is structured in five major steps (fig. 2):

- Definition of the main target (step 1): The target is to develop a XML schema for sustainability reports. The schema has to simultaneously incorporate a variety of issues on sustainability, and the requirements of relevant regulations, standards, guidelines and manuals, especially at the European level.

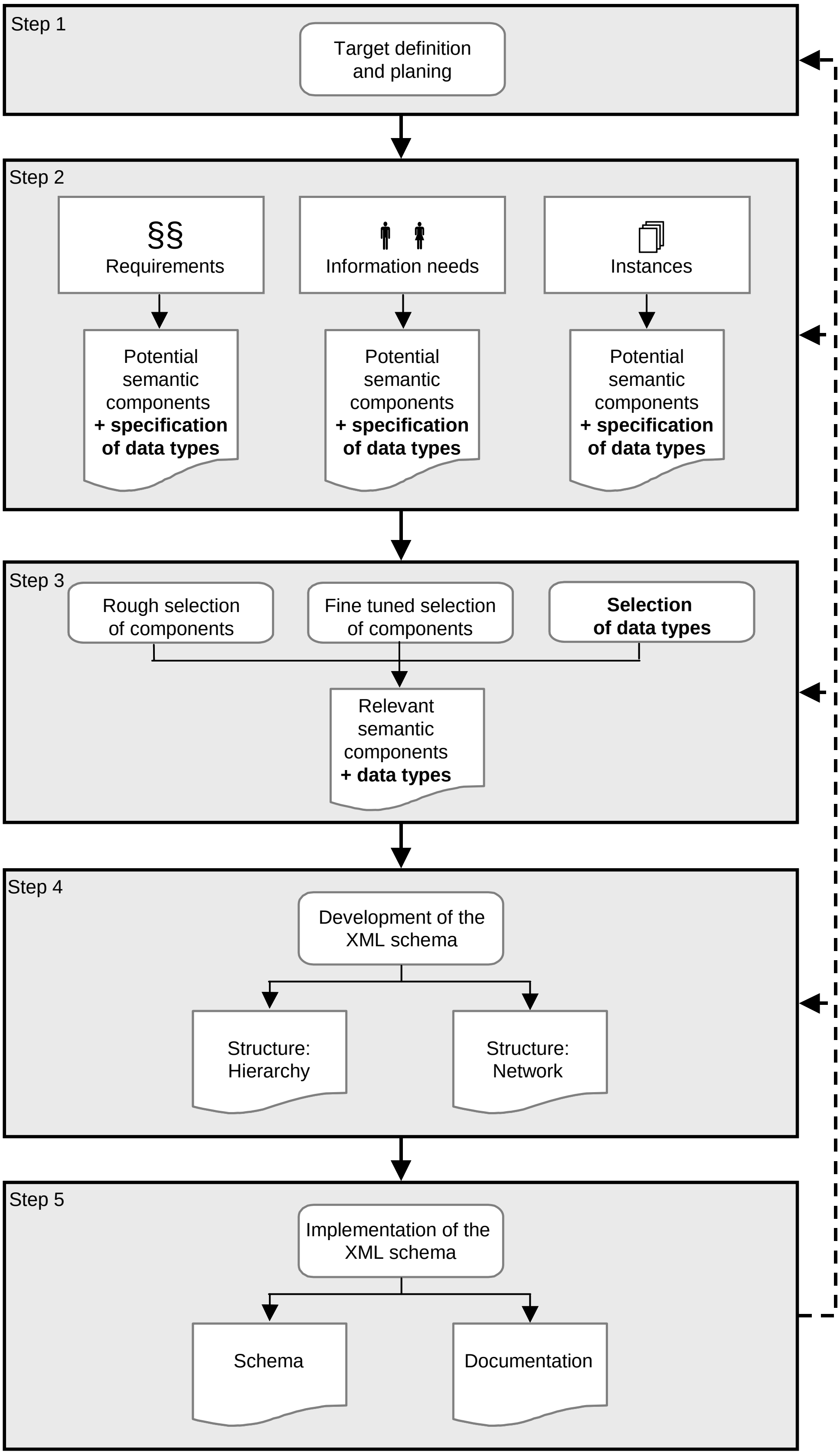


Figure 2: Process model for the development of XML schemas

- Identification of possible semantic components (step 2): According to the target, a multitude of resources need to be analyzed to extract possible contents from relevant regulations, standards, guidelines, already available reports and users' needs and preferences, e.g. the GRI (2002) "sustainability reporting guidelines", the revised European "Eco Management and Audit Scheme" (EC, 2001), the international standard ISO 14001 on "environmental management systems" (DIN, 1996), the German standard DIN 33922 "environmental reports for the public" (DIN, 1997), the early international guideline on "company environmental reporting" (UNEP and SustainAbility, 1994), its German counterpart "environmental reports – environmental statements" (future and IÖW, 1994), and a recently published publicly available specification (PAS) on "data exchange between enterprise resource planning (ERP) systems and environmental information systems" (Lang et al., 2003). This task identifies the pool of possible semantic components the schema may contain. Further, for all resources taken into account, the data types have to be identified and specified.
- Selection of relevant semantic components (step 3): From the pool of possible semantic components, a catalogue of actually relevant contents needs to be developed through a verification procedure. Using the full GRI (2002) content, the result is a total of 395 semantic components (using the full G3 content, the result will decrease to about 260 components). If any sector supplements may be included, then the total amount will increase up to 415-466, depending on the industry sector. All components incorporated are classified as "must be" (required) or "might be" (optional), and on which resource they are rooted. In addition, all data types used for a certain semantic component have to be determined and analyzed through a verification procedure in terms of redundancy. The result is a catalogue of relevant contents specified by certain data types.
- Design of the schema (step 4): Based on the catalogue above, the schema has to be designed. Therefore, all selected components can be organized in a hierarchy, typical for XML documents (fig. 3).

Figure 3: Schema for advanced sustainability reports, illustration

- Implementation of the document type model (step 5): Finally, the schema needs to be implemented, i.e. noted according to XML and transformed into an XML Schema Definition (XSD). Further, a manual needs to be prepared for documentation, further development and regular updates.

Employing an XML schema offers an impressive array of benefits, improves a company's information management, supports its reporting workflow, allocates its resources efficiently, exactly meeting requirements proposed by emerging guidelines, and helps to communicate with its target groups in a meaningful way, that is, providing interactivity, producing tailor-made reports, and facilitating stakeholder dialogue. In total, on the basis of a schema, companies are enabled to provide integrated and customized sustainability reports, prepared by machine processing and generated in an automated manner.

2.4 ICT-architecture

Reaping the media-specific benefits that the internet and XML may carry requires a proper ICT-architecture, suitable to operate XML documents and appropriate to provide single source cross media publishing. For that reason, software tools with at least a three-tier architecture seem to be needed:

- The basic data layer (I) contains several sources where the schema, style sheets, user profiles and a number of other XML documents are stored. These sources include relevant data, metadata and thesauri. The data layer is managed through a database server.
- The application layer (II) contains different services and applications to generate and distribute reports in an automated manner by machine processing. This complex layer is used as a data integrator responsible for system management and is accessed through an application server.
- The presentation layer (III) represents an interactive user interface that is used for submitting users' information needs as well as for presenting reports. The presentation layer provides easy access via a standard internet browser, e.g. Netscape Navigator or Microsoft Internet Explorer.

For example, the procedure of an advanced reporting system is depicted in figure 4.

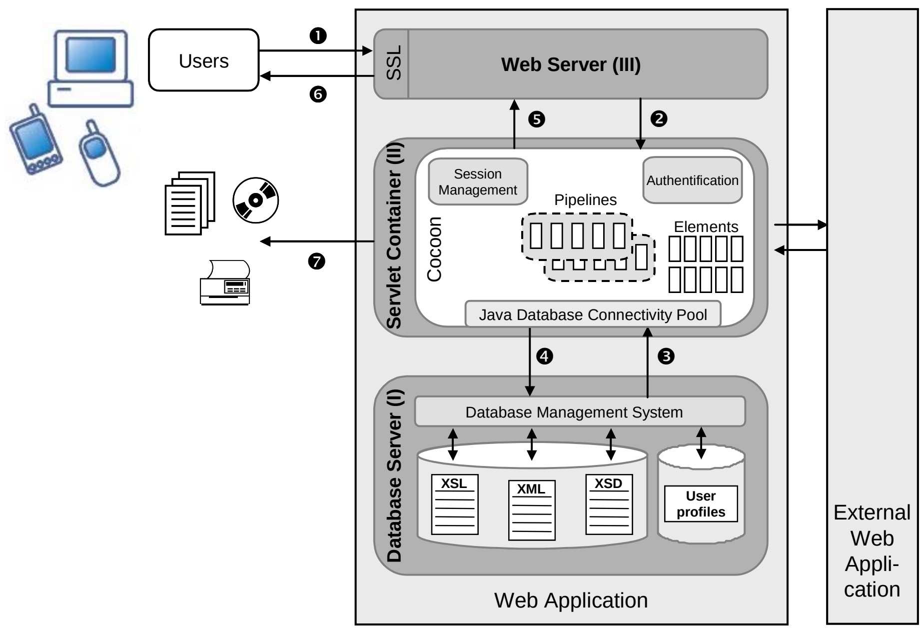


Figure 4: Web application for advanced sustainability reporting, illustration

For demonstration, we developed a set of XSL (eXtensible Stylesheet Language) style sheets (fig. 5), providing a total of nine combinations: This set of style sheets demonstrates how certain contents can be extracted from the same underlying XML-based report, and how these contents might be arranged and presented according to certain user preferences or exactly meeting the requirements of regulations or guidelines in an automated fashion.

<table>
<tr><th>Layout
Content</th><th>Print media</th><th>Computer-based media</th><th>Processing in Environmental Information Systems</th></tr>
<tr><td>Satisfying investors’ needs</td><td rowspan="3">• Plain text
• Table of contents
• PDF</td><td rowspan="3">• Hypertext structure
• Multimedia features
• Navigation
• HTML</td><td rowspan="3">• Plain text
• Marked with tags
• XML</td></tr>
<tr><td>Fulfilling the require-ments of EMAS II</td></tr>
<tr><td>Meeting the GRI guidelines</td></tr>
</table>

Figure 5: Set of different views based on an underlying XML-based sustainability report through specific XSL style sheets

Together, the framework describes a comprehensive and ICT-based reporting approach in a broader sense, including its variety of value added features. Due to its fully supported underlying ICT-infrastructure, an internet-based reporting system of this kind provides a set of important contents, different media, corresponding distribution principles and various presentation styles.

The ICT architecture discussed above has been realized in a software prototype as a practical application (Isenmann et al., 2004, Isenmann, 2008). At the heart of its ICT-architecture lies Cocoon (application layer, II), a Java-based, modular structured, open source publishing software (Apache 2008), able to perform the schema (see fig. 4) and other XML documents, and suitable to provide customized reports on the fly. Today, the prototype can be considered an adaptable reporting system providing individualized reports. The next is the development towards personalization and an adaptive stage. In sum, companies are in a position to make progress in customization step by step, starting with stereotyping and next moving towards a reporting system able to provide individualized or personalized reports. Soon the prototype may be implemented in a number of trail-blazing small and medium sized German enterprises.

3. CONCLUSIONS

Despite its overall usefulness for communicating sustainability issues, online reporting cannot solve a number of current problems discussed in the field (Gray, 2002; Gray and Milne, 2002; Unerman and Bennett, 2004): e.g. should sustainability reporting still be voluntary or must it be mandatory? Is it meaningful to disclose sustainability issues through well established annual and financial reports, or should information be covered just in (standalone) sustainability reports? How could companies’ sustainability performance be conceptualized? What are proper methods of measurement? Which are the most meaningful indicators? Should stakeholders be involved, and what may audit, accounting control, and verification look like? No less importantly, where are the systems boundaries, defining which effects are within the reporting scope and which are not?

Further, online reporting opens up a host of new questions, e.g. with respect to the target groups addressed and those actually reached. Among technical aspects of online communication and matters of efficient information management, a credible effort in sustainability reporting will have to address issues such as the digital divide, restricted access etc. On both ends of the communication link appropriate ICT infrastructure is needed; not just with the companies communicating (reporters), but more importantly with the stakeholders that need to be actually reached (report users). ACCA (2001) compiled some “cardinal sins” of online reporting in order to make its limitations clear.

Looking at the variety of methods firms are adopting for sustainability communication, CSR Europe (2000) point out that reporting belongs to the least focused and coherent means of communication, compared to standards & labels, awards & events, stakeholder consultation, or cause-related marketing. Although an institutional infrastructure around corporate sustainability – resulted in the evolution of initiatives such as the GRI, the social investment movement, and related efforts placing more emphasis on corporate responsibility, accountability and transparency – has emerged, it is evident: sustainability reporting has some way to go before becoming a reliable and trustworthy activity. It is obvious, though, that the aim of providing a "clear and fair view" such as is (supposedly) attained in financial reports is illusory owing to the nature of sustainable development as a social process, where measurable "hard" data will never tell the whole story. From a reporter's perspective, sustainability is a notoriously complex issue to report. From a report user's point of view, this makes performance comparisons, benchmarking, rating and ranking rather difficult.

In contrast to the binary logic of recommending either print media or computer-based media as the only forms of a sustainability report, it is argued here for a cross media reporting approach that relies on an underlying ICT infrastructure. Such an approach may be based on the internet, preferably uses the benefits of XML, and provides powerful support along the whole reporting workflow. Further, such advanced reporting keeps companies in a position to provide sustainability reports and other communication vehicles on a variety of media, based on a single data source that serves as a shared publishing basis.

DiPiazza and Eccles (2002, 127) state that "corporate information, in all its growing quantity and complexity can be – and in reality must be – communicated more effectively with the use of new technology. Reported information needs to break away from the constraints of paper-based formats." Numerous target groups are no longer satisfied solely with reports on print media or mere electronic duplicates. Sustainability reporting is becoming increasingly relevant for decision-making, and responding to multiple inquiries that a variety of stakeholder groups are making to companies is really time-consuming and costly (Axelrod, 2000). Rather than endure these procedures, companies are recognizing the value in having a readily available tool for providing the information needed.

In terms of reporting costs, DiPiazza and Eccles (2002) expect that an advanced reporting approach based on internet technologies and using XML can save companies significant resources, up to 60% of the costs compared with orthodox reporting methods. This cost saving potential might be exploited to a certain extent because orthodox reporting methods and major parts of current practice still rest on an underlying ICT-architecture that is still aimed at print media and that requires manual interfaces and processes impeding the consolidation of computer-based documents. Currently, it is rather time-consuming and labour-intensive to transfer reports initially planned for hard copy into WWW-reports in the proper format for internet use and related applications.

While few commercial software tools are available – like Enablon SD-CSR, SoFi CSM, Credit360, Proventia EnQ CSM, Technidata EH&S Compliance – and others under development (Isenmann, 2008), some pioneers are already employing portals (Moore, 2003; Isenmann and Marx Gómez, 2008) or using web content management systems to support the whole reporting workflow and to improve underlying core processes, from material and process flows to eco input-output-inventory (Marx-Gómez and Rautenstrauch, 2001). Such ICT-systems allows reports' content to be updated with quantitative data, stored, retrieved, edited, controlled and then output cross media in an automated cost saving manner. Technically, such systems typically include databases, workflow tools and editorial applications. It eventually leads to more collaborative work, inside and outside the company.

Pioneering companies have started or will start to implement internet-based applications in the near future. Verie Sandborg, Baxter International's manager of environmental health and safety requirements regards a good environmental or sustainability report as an excellent source for responding to formalized requests for environmental or sustainability information (Axelrod, 2000).

Many of the questions asked are already answered in comprehensive reports. However, it would be helpful to have a fully ICT-based online reporting system: users could extract the information they need from a publishing database, and create an automatically generated customized report themselves, i.e. users are generate their own “reports à la carte”, simply selecting keywords, clicking on preferences on a menu or choosing a certain guideline – perhaps creating a sustainability report in accordance with the GRI-guidelines at one’s fingertips (Isenmann, 2008).

For some companies, advanced sustainability reporting might seem like a nice extra or just a buzzword in comparison to orthodox practice, traditional reporting focused on print media. The unique capabilities and benefits of such reporting approach, however, elevate it beyond the status of a mere “gadget”. Internet technologies and services employed with XML and performed through web content management systems or other sophisticated software tools can do more than just offer new channels for report distribution or presentation.

On the one hand, sustainability reporting using the XBRL taxonomy for sustainability reports is seen a proper information and communication technologies standard. It opens the window for further improvements, especially in terms of efficiencies and automated processes. On the other hand, such harmonized reference architecture fosters standardization in whole field of corporate reporting, be it financial, environmental, sustainability reporting or its certain form as Chemical Safety Reporting. Furthermore, transaction costs while exchanging information between organizations and governmental bodies can be reduced substantially. Finally, the presented XBRL taxonomy promotes the development of smart software tools and the integration of non-financial reporting into common business information systems.

All in all, experience and knowledge gained in SR could or in fact should be transferred to Chemical Safety Reporting and its currently emerging requirements. As Chemical Safety Reporting is mandatory, all stakeholders involved should benefit from developments in the closely related field. For this reason, if no other, the implementation of REACH could be streamlined and efficiencies could surely be exploited.

REFERENCES

Andrew, J. (2003): Corporate governance, the environment, and the internet. Electronic Green Journal 19(December). Online: <http://egj.uidaho.edu/egj19/andrew1.html>, access 2004-04-04.

Apache (2008): Apache Cocoon Project. Online: <http://cocoon.apache.org>, access 2008-03-24.

Arndt, H.-K.; Isenmann, R.; Brosowski, J.; Thiessen, I.; Marx Gómez, J. (2006): Sustainability reporting using the eXtensible Business Reporting Language (XBRL). Tochtermann, K.; Scharl, A. (Eds.). Managing Environmental Knowledge. Aachen: Shaker, 75-82.

Association of Chartered Certified Accountants (ACCA) (1998): Making values count: contemporary experience in social and ethical accounting, auditing, and reporting. London: The Certified Accountants Educational Trust.

Association of Chartered Certified Accountants (ACCA) (2001) Environmental, social and sustainability reporting on the world wide web: A guide to best practice. London: The Certified Accountants Educational Trust.

Association of Chartered Certified Accountants (ACCA) (2004): Towards transparency: Progress on global sustainability reporting 2004. London: Certified Accountants Educational Trust.

Australian Government Department of the Environment and Heritage (2003): Corporate sustainability – an investor perspective. The Mays Report Canberra, Environmental Protection Branch, Department of the Environment and Heritage. Online: <http://deh.gov.au/industry/finance/publications/index.html>, access 2004-03-22.

Axelrod, R.A. (2000): Brave new words. The financial value of environmental communications. Environmental Quality Management 9(4): 1-11.

Beattie, V.; Pratt, K. (2003): Issues concerning web-based business reporting: An analysis of the views of interested parties. The British Accounting Review 35(2): 155-187.

Brophy, M.; Starkey, R. (1996): Environmental reporting. R. Welford (Ed.). Corporate environmental management. Systems and strategies. London: Earthscan, 177-198.

Brosowski, J.; Isenmann, R.; Beisel, M.; Marx Gómez, J. (2004): XML schema for sustainbility reports meeting the needs of the GRI guidelines. Minier, P.; Susini, A. (Eds.). Sh@ring. Proceedings EnviroInfo 2004. 18th International Conference Informatics for Environmental Protection. Geneve: Editions du Tricorne, 184-193.

Business for Social Responsibility (BSR) (2008): Environmental, social and governance: Moving to mainstream investing? Online: <http://bsr.org>, access 2008-08-28.

Clarke, T. (2001): Balancing the triple bottom line: Financial, social and environmental performance. Journal of General Management 26(4): 16-27.

Clausen, J.; Loew T, Klaffke K, Raupach M, Schoenheit I. (2001): The INEM sustainability reporting guide – a manual on practical and convincing communication for future-oriented companies. Hamburg: International Network for Environmental Management (INEM). Online: <http://www.inem.org/free_downloads>, access 2002-03-22.

CSR Europe (2000): Communicating corporate social responsibility. Brussels (Belgium): CSR Europe.

Dalen van, M. (1997): Company environmental reporting. Conditions for the optimal information structure of environmental reports. Dissertation thesis. University of Humberside (UK) and Zeeland Polytechnic Flushing (The Netherlands).

Danish Commerce and Companies Agency (DCCA) (2006): People & Profit. A practical guide to corporate social responsibility. Online: <http://www.eogs.dk/sw28291.asp>, access 2008-09-22.

Deloitte Touche Tohmatsu International (DTTI), International Institute for Sustainable Development (IISD), SustainAbility Ltd. (1993): Coming clean - corporate environmental reporting, opening up for sustainable development. London (UK): DTTI.

DiPiazza, S.A.; Eccles, R.G. (2002): Building corporate trust. The future of corporate reporting. New York: Wiley.

Elkington J. (1997): Cannibals with forks: The triple bottom line of 21st century business. Oxford: Capstone.

European Commission (2008): Commission regulation (EC) No 987/2008 of 8 October 2008 amending Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) as regards Annexes IV and V. Official Journal of the European Union L 268/14-L 268/19.

European Commission (EC) (2002): European SMEs and Social and Environmental Responsibility. Observatory of European SME 2002/No. 4, prepared by KPMG Special Services and EIM Business & Policy Research in the Netherlands in co-operation with European Network for SME Research (ENSR), and Intomart Luxembourg, Office for official publications of the European Communities.

European Communities (EC) (2001): Regulation (EC) No. 761/2001 of the European Parliament and of the Council of 19 March 2001 allowing voluntary participation by organisations in a Community eco-management and audit scheme (EMAS). Official Journal of the European Communities, L 114.

Fichter, K. (1998): Umweltkommunikation und Wettbewerbsfähigkeit. Wettbewerbs-theorien im Lichte empirischer Ergebnisse zur Umweltberichterstattung von Unternehmen [Environmental communication and competitiveness. Concepts of competition in the light of empirical results on corporate environmental reporting]. Marburg: Metropolis.

Figge, F.; Schaltegger, S. (2000): What is “stakeholder value”? Developing a catchphrase into a benchmarking tool. Lüneburg, University of Lüneburg and UNEP.

Futerra Sustainability Communications (2008): The greenwash guide. Online: <www.futerra.co.uk>, access 2008-09-17.

future eV, Institute for Ecological Economy Research (IÖW) (1994): Environmental reports – environmental statements. Guidelines on preparation and distribution. Osnabrück: future eV.

German Institute for Standardisation eV (DIN) (1996): Environmental management systems – specifications with guidance for use. Berlin (Germany): Beuth.

German Institute for Standardisation eV (DIN) (1997): DIN 33922. Environmental reports for the public. Berlin (Germany): Beuth.

Global Reporting Initiative (GRI) (2002): Sustainability reporting guidelines. <http://www.globalreporting.org>, access 2002-10-15.

Global Reporting Initiative (GRI) (2006): Sustainability reporting guidelines. Version 3.0. Amsterdam: GRI.

Gray, R. (2002): Social and environmental responsibility, sustainability and accountability: Can the corporate sector deliver? The Corporate Citizen 2(3): 9-16.

Gray, R.; Milne, M. (2002): Sustainability reporting: Who's kidding whom? Chartered Accountants Journal of New Zealand 81(6): 66-70.

Hamscher, W. (Ed.) (2007): Financial reporting taxonomies architecture 1.0. Recommendation dated 2005-04-25. Online: <http://www.xbrl.org/technical/guidance/FRTA-RECOMMENDATION-2005-04-25.htm>, access 2007-12-14.

Hesse, A. (2007): Sustained added value. Information of investors and analysts for sector-specific "Sustainable Development Key Performance Indicators" (SD-KPIs) in management commentaries in German companies. Study prepared for Deloitte Touche Tomatsu. Online: <http://deloitte.com/de>, issued 1/2007, access: 2008-06-21.

Hund, G.; Engel-Cox, J.; Fowler, K. (2004): A communications guide for sustainable development. How interested parties become partners. Columbus: Battelle.

Interbrand (2008): Best global brands 2008. Creating and managing brand value. Online: <http://www.interbrands.com>, access: 2008-09-29.

International Institute for Sustainable Development (IISD), Deloitte & Touche, Business Council for Sustainable Development. 1992. Business strategy for sustainable development. Leadership and accountability for the 90'. Winnipeg: IISD.

Isenmann, R. (2004): Internet-based sustainability reporting. International Journal of Environment and Sustainable Development 3(2): 145-167.

Isenmann, R. (2005): Corporate sustainability reporting – a case for the internet. Hilty L.M.; Seifert, E.K.; Treibert, R. (Eds.). Information systems for sustainable development. Hershey: Idea Group, 164-212.

Isenmann, R. (2008): Environmental management information systems – illustrations from online communication and sustainability peporting. Proceedings of the iEMSs Fourth Biennal Meeting. International Congress on Environmental Modelling and Software (iEMSs 2008). Sánchez-Marrè, M.; Béjar, J.; Comas, J.; Rizzoli, A.E.; Guariso, G. (Eds.). International Environmental Modelling and Software Society, Barcelona, Catalonia, July 2008 vol. 3, 1636-1644.

Isenmann, R.; Arndt, H.-K.; Bey, C.; Marx Gómez, J.; Solsbach, A. (2008): XBRL-Taxonomie für Nachhaltigkeitsberichte – Vorschlag einer harmonisierten Dokument-struktur zur Online-Kommunikation von Organisationen [XBRL taxonomy for sustainability reports – proposal of a harmonized document structure for online communication at corporate level]. Bichler, M.; Hess, T.; Krcmar, H.; Lechner, U.; Matthes, F.; Picot, A.; Speitkamp, B.; Wolf, P. (Eds.). Multikonferenz Wirtschaftsinformatik 2008. Berlin (Germany): GITO, 601-613.

Isenmann, R.; Beisel, M.; Brosowski, J.; Marx Gómez, J. (2004): Software tool for single source multiple media sustainability reporting in: Minier P, Susini A. Sh@ring. Proceedings EnviroInfo 2004, 18th International Conference Informatics for Environmental Protection. Geneve: Editions du Tricorne, 152-162.

Isenmann, R.; Bey, C.; Welter, M. (2007): Online reporting for sustainability issues. Business Strategy and the Environment 16(7): 487-501.

Isenmann, R.; Brosowski, J.; Lenz, C. (2003): Target group tailored corporate environmental reporting – stereotyping, individualisation, personalisation. Gnauck, A.; Heinrich, R. (Eds.). The information society and enlargement of the European Union. Proceedings EnviroInfo 2003, 17th International Conference Informatics for Environmental Protection. Marburg (Germany): Metropolis, 533-540.

Isenmann, R.; Kim, K. (2006): Interactive sustainability reporting – developing towards more target group tailoring and stakeholder dialogue. Schaltegger, S.; Bennett, M.; Rikhardsson, P. (Eds.). Sustainability accounting and reporting. Amsterdam: Kluwer, 533-555.

Isenmann, R.; Lenz, C. (2001): Customized corporate environmental reporting by internet-based push- and pull-technologies. Eco-Management and Auditing 8(2): 100-110.

Isenmann, R.; Lenz, C. (2002): Internet use for corporate environmental reporting: Current challenges – technical benefits – practical guidance. Business Strategy and the Environment 11(3): 181-202.

Isenmann, R.; Marx Gómez, J. (2004): How to provide customized environmental reports properly. Scharl A. (Ed.). Environmental online communication. London: Springer, 173-182.

Isenmann, R.; Marx Gómez, J. (Eds.) (2008): Internetbasierte Nachhaltigkeitsbericht-erstattung. Maßgeschneiderte Stakeholder-Kommunikation mit IT [Internet-based sustainability reporting. Tailored stakeholder communication with IT] Berlin (Germany): Erich Schmidt.

Jensen, R.E.; Xiao, J.Z. (2001): Customized financial reporting, networked databases, and distributed file sharing. Accounting Horizons 15(3): 209-222.

Kolk, A. (2004): A decade of sustainability reporting: developments and significance International Journal of Environment and Sustainable Development 3(1): 51–64.

KPMG (2000): Beyond the numbers: How leading organizations are linking values with value to gain competitive advantage. KPMG’s Assurance & Advisory Services Center (AASC), KPMG.

KPMG (2008): International survey of corporate responsibility reporting 2008. Amstelveen: KPMG Sustainability Services.

Lang, C.; Rey, U.; Wohlgemuth, V.; Genz, S.; Pawlytsch, S. (2003): Data exchange between ERP-systems and environmental information systems. Publicly available specification (PAS) 1025. Berlin (Germany): Beuth.

Larsen, L.B. (2000): Strategic implication of environmental reporting. Corporate Environmental Strategy 7(3): 276-287.

Lenz, C. (2003): Empfängerorientierte Unternehmenskommunikation – Einsatz der Internet-Technologie am Beispiel der Umweltberichterstattung [Customised environmental communication – use of internet illustrated through corporate environmental reporting]. Köln: Eul.

Line, M.; Hawley, H.; Krut, R. (2002): The development of global environmental and social reporting. Corporate Environmental Strategy 9(1): 69-78.

Lippman, S. (2001): Supply chain environmental management. Environmental Quality Management 11(2): 11-14.

Lodhia, S. (2004): Corporate environmental reporting media: A case for the world wide web. Electronic Green Journal, 20 (spring). Online: <http://egj.lib.uidaho.edu/egj20/lodhia1.html>, access 2004-11-23.

Marshall, S.R.; Brown, D. (2003): Corporate environmental reporting: What’s in a metric? Business Strategy and the Environment 12(2): 87-106.

Marx Gómez, J.; Rautenstrauch, C. (Eds.) (2001): Von der Ökobilanzierung bis zur automatisierten Umweltberichterstattung mit Stoffstrommanagementsystemen – eine Fallstudie [From eco input-output-inventory towards automated environmental reporting with material flow management systems – a case study]. Aachen: Shaker.

Mesterharm, M. (2001): Integrierte Umweltkommunikation von Unternehmen. Theoretische Grundlagen und empirische Analyse der Umweltkomunikation am Beispiel der Automobilindustrie [Integrated environmental communication at corporate level. Conceptual foundation and empirical analysis, illustrated by the automotive sector]. Marburg Metropolis.

Moore, M.A. (2003): Through the eye of a needle. The importance of portals to environmental information management. Environmental Quality Management 13(1): 55-66.

Pleon Kohtes & Klewes (2005): Accounting for good: The global stakeholder report 2005. Pleon: Amsterdam (The Netherlands), Bonn (Germany).

Raar, J. (2002): Environmental initiatives: Towards triple-bottom line reporting. Corporate Communications 7(3): 169-183.

Rikhardsson, P.; Andersen, A.J.R.; Bang, H. (2002) Sustainability reporting on the internet. A study of the Global Fortune 500. Greener Management International (40): 57-75.

Schraml, T. (1997): Operationalisierung der ökologieorientierten Berichterstattung aus Sicht des Informationsmanagements [Specification of environmental reporting from an information management perspective]. Dissertation. Technische Universität Dresden.

Scott, P.; Jackson, R. (2002): Environmental, social and sustainability reporting on the web. Best practices. Corporate Environmental Strategy 9(2): 193-202.

Shepherd, K.; Abkowitz, M.; Cohen, M.A. (2001): Online corporate environmental reporting: Improvements and innovation to enhance stakeholder value. Corporate Environmental Strategy 8(4): 307-315.

Sociaal Economische Raad (SER) (2001): Corporate social responsibility. A Dutch approach. Assen, Koninklijke.

SocialFunds (Socially responsible investing information concerning corporations, mutual funds, and more) (2000): Investors prod CEO's to adopt new reporting standards. <http://socialfunds.com/news/article>, © 2000, access 2000-12-16.

SustainAbility Ltd. and United Nations Environment Programme (UNEP) (1999) Engaging stakeholders 1999. The internet reporting report. London: Beacon.

Teo, H.-H.; Oh, L.B.; Liu C, Wei, K.K. (2003): An empirical study of the effects of interactivity on web user attitude. International Journal of Human-Computer Studies 58(3): 281-305.

Unerman, J.; Bennett, M. (2004): Increased stakeholder dialogue and the internet: Towards greater corporate accountability or reinforcing capitalist hegemony? Accounting, Organziations and Society 29(7): 685-707.

United Nations Environment Programme Industry and Environment (UNEP), SustainAbility (1994): Company environmental reporting. A measure of the progress of business & industry towards sustainable development. Technical report 24. Paris: UNEP.

Visser, W. (2002): Sustainability reporting in South Africa. Corporate Environmental Strategy 9(1): 79-85.

Volkswagen AG (2003): Mobility and responsibility. Improving social performance reporting. Online: <http://www.vw-ag.de/english/defaultIE.html>, access 2003-04-04.

Waddock, Sandra (2008): Building a new institutional infrastructure for corporate responsibility. Academy of Management Perspective 22(3): 87-108.

Wheeler, D.; Elkington, J. (2001) The end of the corporate environmental report? Or the advent of cybernetic sustainability reporting and communication. Business Strategy and the Environment 10(1): 1-14.

World Business Council for Sustainable Development (WBCSD) (2002): Communications and stakeholder involvement guidebook for cement facilities. Report prepared by the Battelle Memorial Institute (Ohio) and Environmental Resources Management (London). Online: <http://www.wbcsdcement.org/pdf/final_report1_2.pdf>, access 2003-09-18.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pilmmann, I.Holoubek, T. Banholz (eds.)
Masaryk University, 2009

eParticipation in the service of environmental democracy: Introducing the U@marenostrum project

E.Karamagioli[a] **and M.P Ortega**[b]

[a] *Gov2u, Anavriton 20, Athens, PC 15124 Greece (evika@gov2u.org)*

[b]*GFI Benelux Offices, Leuven, B-3001 Belgium (mportegar@gfi.es)*

Abstract: Environmental policy is one area where there is a great deal of public support for action at a Europe-wide level. Public involvement can improve the quality of decision-making, raise citizen awareness of environmental issues and increase public understanding of the decision-making process. Public involvement can be maximized by Internet-based approaches and the Web should be seen as a means of enhancing current practices. Geographic information and Internet technologies used for eParticipation may be employed as positive tools for increasing the transparency of decisions on environmental issues. The effective engagement of citizens by governments rests on their recognition of access to environmental information as a basic precondition, consultation as central to policy-making and public participation as a relationship based on partnership. The new tools offered by ICTs can offer assistance in each of these domains. Their impact can also be greatly enhanced through use in combination with traditional, "offline" methods.
The proposed paper will introduce the objective and expected results of an eParticipation project addressing the aforementioned issue of environmental democracy. The pilot project entitled "U@MARENOSTRUM" aims to involve, thanks to ICT and GIS, citizens and local actors (associations) from the Mediterranean coastal zones in decision-making processes for the adoption and implementation of water and marine environmental protection policies and legislation in the Mediterranean region in accordance with EU environmental legislation.

Keywords: eParticipation; ICT; GIS; environmental democracy

1. INTRODUCTION

Environmental policy is one area where there is a great deal of public support for action at a Europe-wide level. The 6th Environment Action Program of the European Community 2002-2012[1] is a decision of the European Parliament and the Council adopted on 22nd July 2002. It sets out the framework for environmental policy-making in the European Union for the period 2002-2012 and outlines actions that need to be taken to achieve them. The 6th EAP stresses the importance of providing adequate environmental information and effective opportunities for public participation in environmental decision-making; thereby increasing accountability and transparency of decision-making and contributing to public awareness and support for the decisions taken.

The effective engagement of citizens by governments rests on their recognition of access to environmental information as a basic precondition, consultation as central to policy-making and public participation as a relationship based on partnership. The new tools offered by

[1] http://ec.europa.eu/environment/newprg/index.htm

ICTs can offer assistance in each of these domains. Their impact can also be greatly enhanced through use in combination with traditional, “offline” methods.

The proposed paper will introduce the objective and expected results of an eParticipation project addressing the aforementioned issue of environmental democracy. The pilot project entitled “U@MARENOSTRUM” aims to involve, thanks to ICT and GIS, citizens and local actors (associations) from the Mediterranean coastal zones in decision-making processes for the adoption and implementation of water and marine environmental protection policies and legislation in the Mediterranean region in accordance with EU environmental legislation.

The project intends to advance the idea and use of eParticipation in environmental decision-making by empowering citizens and local associations to use the state-of-the-art in order to debate and engage with the implementation of EU environmental legislation. By providing a specific solution that will involve actors from both citizens and public institutions, the scope of the project is to introduce a bottom-up relationship between decision-makers and citizens, to confront the perceived democratic deficit, to connect citizens with politics and finally to simplify the complexity of decision-making processes.

2. DEFINING EPARTICIPATION

Faced with declining citizen participation in democratic processes across Europe in recent years and growing demands from citizens and other stakeholders to have a more influential role in policymaking, the EU has been pioneering research in methods that empower citizens by making their voices heard, both through official and informal forums. A wide range of relevant ICT tools have emerged as a result, the most common of which are net-voting and feedback mechanisms that allow the public to participate more substantively in advocacy. They are the modern implementations of traditional means of democracy. eParticipation can be defined as the use of information and communication technologies (ICT) to broaden and deepen political participation by enabling citizens to connect with one another and with their elected representatives. eParticipation tools consist of software applications, products, tools and components that are used in the domain of participation.

eParticipation is considered as having a number of advantages over the conventional participation methods because:

- The information about the issues being discussed is available from any location that has Web access 24 hours a day and 7 days a week. This gives the opportunity for more people to participate in public consultations.
- The participation is not restricted by geographical location like in the case of the meetings in public places.
- Participation is relatively anonymous and less confrontational as compared to a face-to-face meeting. This may encourage the silent majority to participate as long as this methods allow practical non- threatening modes of interaction by being anonymous.
- It allows the sharing and exchange of information and ideas in an effective manner.

The past few decades, a large volume of literature has been published dealing with various aspects of hydrogeology and salinity in atoll island aquifer systems and its numerical modelling.

3. ELECTRONIC PARTICIPATION IN THE SERVICE OF ENVIRONMENT

3.1 General Overview

As Information and Communications Technologies (ICTs) advance, new participation methods based around these technologies have been developed making online participation possible. An example of these technologies is the GIS (geographic information system) which has the capacity to integrate information from a variety of sources into a spatial context and is well suited to support decision making procedures. It is an information system for capturing, storing, analyzing, managing and presenting data which are spatially referenced. GIS can act as a tool in helping the decision-makers evaluate alternatives, visualise choices and explore certain alternatives.

GIS and the Web hold a great potential for public use, allowing wider involvement in environmental decision making.The idea of a web-based Public Participation GIS revolves around the concept of public participation in the use of spatial data leading to increased community involvement in policy-setting and decision-making. It composes peoples' spatial knowledge in the forms of virtual or physical, 2 or 3 dimensional maps used as interactive vehicles for discussion, information exchange, analysis and as support in advocacy. This visual aspect of GIS technology makes it a key ingredient in raising public awareness about environmental issues. The ability to manage, analyze, and delineate a variety of geospatial data makes GIS essential for targeting areas in need and communicating ways to solve the environmental problems related to that areas. Using a `dynamic map' that is interactive and provides particular pieces information about features on it, allows the user to elicit greater detail about issues and problems in hand such as the relative location of features and proposed developments.

3.2 The Aarhus Convention

Public participation in environmental decision making has become in the last decades the goal of different recommendations and declarations at the international level, elaborated by the European Union. In June 1998, the Convention on Access to Information, Public Participation in Decision-making and Access to Justice in Environmental Matters in Aarhus, Denmark, also called the Aarhus Convention, was signed and ratified by the European Union and countries in central Asia. The Århus Convention entitles the public - individuals and environmental organisations - to participate in environmental decision-making. This means that public authorities have to inform them in time of relevant proposals, so that they can submit their comments. These comments need to be taken into consideration. The authorities are also required to inform the public of their final decisions and the underlying reasons.

Under the Århus Convention, public authorities are obliged to organize, make accessible and disseminate the environmental information in their possession. This means they are obliged to keep and share the environmental information relevant to their functions, such as the quality of water, air and the atmosphere, land and biological diversity; energy, noise, development plans and policies; and the effects of these on human health, safety and the environment. This information has to be made accessible and disseminated, for example by putting it on the Internet, informing the media or using other channels of communication. Environmental democracy is subject of the Aarhus Convention that goes to the heart of the relationship between people and governments. The Convention is not only an environmental agreement; it is also a Convention about government accountability, transparency and responsiveness.

On 26 May 2003 the Directive 2003/35/EC[1] for public participation was adopted in respect of the drawing up of certain plans and programs relating to the environment and amending with regard to public participation and access to justice. Provisions for public participation in environmental decision-making are furthermore to be found in a number of other environmental directives, such as Directive 2001/42/EC[2] of 27 June 2001 on the assessment of certain plans and programs on the environment and Directive 2000/60/EC[3] of 23 October 2000 that establishes a framework for Community action in the field of water policy. The latter was adopted on 6 September 2006: Regulation (EC) N° 1367/2006[4] on the application of the provisions of the Aarhus Convention on Access to Information, Public Participation

[1] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32003L0035:EN:NOT
[2] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32001L0042:EN:NOT
[3] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32000L0060:EN:NOT
[4] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32006R1367:EN:NOT

in Decision-making and Access to Justice in Environmental Matters to Community institutions and bodies and became of application on 17 July 2007**.**

4. THE U@MARENOSTRUM PROJECT

4.1 The project overview

The pilot project entitled "U@MARENOSTRUM" aims to involve, thanks to ICT and GIS, citizens and local actors (association) from the Mediterranean coastal zones in decision-making processes for the adoption and implementation of water and marine environmental protection policies and legislations in the Mediterranean region in accordance with the EU environmental legislation.

The project intends to advance the idea and use of eParticipation in environmental decision-making by empowering the citizens and local associations to use the state-of-the-art in order to debate and engage with the implementation of EU environmental legislation. By providing a specific solution that will involve actors from both citizens and public institutions, the scope of the project is to introduce a bottom up relationship between decision-makers and citizens, to confront the perceived democratic deficit, to connect citizens with politics and finally to simplify the complexity of decision-making processes.

The proposed platform addresses two of the pillars of Aarhus Convention, namely the access to information and the public participation in decision-making. The environment and more specifically water management and marine environment protection were chosen for the proposed trial project since they are emerging EU policy issues that have an impact on all countries within the European Union. The environment affects the daily lives of all citizens. That is why citizens are expected to participate in the discussions of the platform. Water is the most basic and critical component of all aspects of human life. Freshwater resources are an essential component of the earth's hydrosphere and an indispensable part of all terrestrial ecosystems. Also, transboundary water resources and their use are of great importance to riparian States. Similarly, the marine water environment - including the oceans and all seas and adjacent coastal areas - forms an integrated whole that is an essential component of the global life-support system. The marine ecosystems are a positive asset that presents opportunities for sustainable development.

Trial applications including the validation of the project will be conducted in 3 end-user locations around the Mediterranean region, namely the Region on Ionian islands (Greece), the French Riviera (France) and the Region of Valencia (Spain). Specific committees from those regional administrative bodies will be assigned to use the platform for a period of 14 months. Three platforms are envisaged. They will be hosted in the participating regions and will be given a regional 'look and feel' (including language) to ensure user participation. The platforms will be specifically designed to serve the citizens of the participating regions. Of course, citizens from other regions will be welcome to use the platforms. The end-users of the platform will be citizens and decision-makers from France, Spain and Greece.

The project will support the citizens and local actors to identify environmental problems that need to be solved immediately in order to result in positive development in the French Riviera, Valencia and Ionian Islands regions as well as providing a way to solve them. At the same time, citizens will be better informed about the EU water and marine environment policies as well as about their national water and marine environmental policies.

Consultations can be developed in order to have the public opinion for developing national maritime strategies or river management basin plans. Furthermore, the project will create a framework for exchange and cooperation between Member States. In this way, the actions in the environmental sector will be made more coherent and the conservation of the marine ecosystems and sustainable water resource management will be supported. The exchange of information will promote cross-border cooperation as well as mutual understanding between the countries involved in the project. Moreover, the project will support stakeholders from the French Riviera, Valencia and Ionian regions to consult each other on all major plans and strategies under consideration that are likely to have a significant adverse environmental impact across boundaries. At the Ionian region cooperation will be promoted between citizens and representatives of the local public authorities in developing environmental plans and actions in the region.

During the trail period an editorial team will be in charge of the publication of the information on the website and of conceiving summaries of the discussions within the participative processes. These summaries will be forwarded to relevant decision-makers from local, regional, national and EU level in order to ensure that the legislative input created by participating citizens really reaches its addressees.

We assume that other European regions will follow the set example and accept and adapt the horizontal U@MareNostrum approach to their region. In this way the project is aiming a high degree of transferability, one of the main objectives of the eParticipation program.

4.2 Objectives

The overall objectives of the project are:

- Reconnect citizens with politics and policy making by strengthening and broadening citizens participation in decision making process
- Address a key issue at EU and national and regional levels too: Environment
- Complement and advance the eParticipation activities already undertaken by the European Union in order to identify what the future orientation of eParticipation
- Provide both decision-makers and the public with clear and understandable information on the decision-making process for the environmental policy field
- Familiarize citizens with the decision making process, making easy to citizens to visualize and identify the stages of decision making process
- Enable people with the same interests to form groups which transverse geographic barriers, and help these groups to participate in decision-making on environmental issues in the Mediterranean region
- Address with the use of ICT multilingual and multicultural problems of people wanted to be involved in the decision making process

4.3 Expected end-users

The expected end-users of U@MARENOSTRUM platform are:

- Citizens from France, Greece and Spain
- Parliamentarians from national assemblies of France, Greece and Spain with EU-policy focus
- Elected representatives from the region of Valencia
- Members of the associations specialized in sectors concerned the environment (protection of the environment ; protection of natural heritage)
- Representatives of research institutes specialized in the sector concerned (water, environment, pollution)
- Port authorities and users of the ports
- Members of organizations for the tourism promotion and territorial planning
- Representatives of certification and control organisms
- Representatives from local administrative bodies

- Members of the European Parliament/ Committee of the Regions linked to French Riviera, Valencia and Ionian Islands regions
- Representatives of planning agencies, development agencies
- Representatives of agencies and institutes specialized in the sectors concerned (planning, environmental protection, coastal protection, natural heritage management, water management)
- Representative of enterprises and private agencies in charge of the sector concerned (water management and water distribution
- Representatives of the private sector
- Members of the Institutions representing farmers and rural sector
- Ministries, regional and local bodies with responsibilities in environmental issues.

4.4 The consortium

The project will be developed during 2 years by a consortium consisting of the following entities:

- GFI – Benelux
- Government to you (Gov2u) – Greece
- Hellenic Centre for Marine Research, (HCMR) – Greece
- Foundation Comunidad Valenciana-Región Europea – Spain
- The Community of French Riviera – France
- Directorate General for Modernisation of Valencia– Spain
- The Region of Ionian Islands – Greece

These members of the consortium are NGOs, research centers and public authorities with rich experience in developing ICT solutions to improve traditional democratic systems and in environmental sector.

5. THE PROPOSED SOLUTION

5.1. Functionalities

The project will lead the development of a portal to capture information on a geographical map. The portal will be very user friendly and easy to use especially by people with limited knowledge of the subject area and low level of IT skills.

The proposed technological platform will be a web-based Public Participation GIS created through the customization of Gov2Demoss, an open source platform that will be integrated with a geographic information system (GIS).The Gov2DemOSS platform is a citizen participation platform which brings forth the latest technology tools to provide a new mechanism of interaction for citizens with their local authorities. It provides an efficient channel for institutions and organizations to keep their communities informed, manage their information repositories, gauge public opinion, interact directly with their constituents, and most importantly, to involve them in the decision making process. Main functionalities include forums, blogs, uploading of documents, online petitions, open letters to elected officials and mayors, messaging and notifications, consultation of Council Plenary session calendar, agendas, documents and newsletters. Gov2DemoSS was awarded the EU Good Practice Label in 2006 and is being used by more than 70 local authorities in Europe.

In addition, the design of the system revolves around Geo Tools which is an open source (LGPL) Java code library which provides standards compliant methods for the manipulation of geospatial data in order to implement Geographic Information Systems (GIS). This GIS will allow the user to view a map, perform zoom and pan operations to assist in visualization and navigation, ask questions such as "what is this building?" and "what is this road?" and then make suggestions about specific features identified from the map .

An important tool of the platform will be “My Agora” that will be used for the creation and management of participative processes on topics related to current policies and also proposals for policies and action plans at transnational, national, local level, which can have associated documents, links, forums for debates, questionnaires and calendar of events related to the participative processes. The documents linked with each participative process could be:

- The implementation of the laws that are already in force
- Draft laws, draft action plans

The platform will also feature a section that visualizes the different stages of the legislative process in a graphical matrix - from the introduction of a new topic, to the consultation process and the discussion and the passing in the local or regional Council or in the Parliament. The graphic matrix will indicate at which stage each selected focus issue currently is. The graphical representation will also include interactive elements which allow the user to learn more about each phase, link to relevant policy documents, and sketch out possible/actual points of discussion/interest. The different steps of the legislative process will be highlighted in different colors depending on their status. The active step will be marked in "green" while all other steps will be marked in "grey" as this would show that they are "inactive". For example, if a proposal is not accepted ("Act is not adopted"), the active box should be shown in red to point out that this proposal has finished.

In order to contact and directly interact with their elected representatives and decision-makers, citizens will be able to use online petitions in which they can create their own petition to express their views, concerns and questions related to environmental issues or to join the petitions created by other citizens on these issues. At the same time, with the support of letter to officials’ tool, each citizen will be able to contact and directly interact with an elected representative or decision-maker from EU, national, regional and local level, as well as reading other citizens’ letters and the answers received by these officials.

A Survey tool will be used to request citizens’ feedback on specific environmental issues and also to measure user satisfaction and the impact of the trails on the users.

The citizens will be encouraged to participate in the discussion forums on specific policy issues of common public interest that will be proposed by the local or national authorities and that could be related with the preparation, modification or review of plans or programs relating to water protection and management. At the same time, citizens will be encouraged to create in the online forums their own topics of discussion related to the protection and conservation of marine environment, water management and planning in the Mediterranean region.

A typical example of use of the platform would be public consultations on draft river basin management plan including program of measures, definition of a zoning environmental protection, elaboration of director plan for development and management of water, law around water aquatic environment, marine strategies, environment legislation that is essential for sustainable water protection. In this framework the citizens would have access using a website to multi-modality channel in terms of response: text entry, seizure zoning on a map with automatic adjustment for the rest of the map using constraints (legal, environmental and so on.). They will be able to register their views about particular issues by placing flags with written comments to appropriate locations on the map and to view and response to the comments placed by other citizens through the flags on the map. An end-user will be able to view a map of the area he is interested in, perform zoom and pan operations to assist in visualization and navigation, ask questions such as "what is this building?" and "what is this road?" (spatial query) and then make suggestions about specific features identified from the map (attribute input).

Additionally in order to increase the number of users we propose to carry out integration with the WEB 3.0 social networking tools like FaceBook, MySpace and Twitter. Public administrations will be able to create their own groups or join groups that are related to topics which they find interesting. Many of these groups attract a large number of citizens that are very active in the corresponding topic. Their integration with the U@MareNostrum will allow information published on public administrations' eParticipation portals to appear on the FaceBook, MySpace, Twittter groups mentioned above.

Within the project a spatial database will be developed that will support the access to geographical information resources. This spatial database will consist in different types of information:

- Legal - concerned with the EU policy and national and local policies for water protection and management within the Mediterranean region, more specific the Valencia, French Riviera and Ionian Islands regions costal zones.
- Environmental data related to the water protection and management in the Mediterranean costal zone.

The information on the portal will be in 3 languages: French, Spanish and Greek. The automatic translation of citizen opinions can be achieved by using BabelFish Web Service.

5.2. Expected Results

Until the end of the project we have set the following objectives to be fulfilled:

1. To advance the concept of environmental democracy via the use of ICT
2. To present new ways of effectively presenting information and communicating addint to the improving the transparency in tracking legislation and decision making process
3. To introduce and test Public Participation GIS to manage geospatial boundary and statistical information that can support well-informed decisions for the management of water and marine environmental protection in the Mediterranean region.
4. To make EU environmental legislation accessible for everyone through electronic means.
5. To support the citizens and local actors to monitor the implementation of the EU water and maritime environmental policies in the in the Mediterranean region.
6. To obtain public input for designing long term environmental policies to be implemented in the French Riviera, Valencia and Ionian Islands regions, which are in line with the EU environmental policies.
7. To make possible for citizens with limited knowledge of the subject area and low level IT skills to profit for eParticipation tools

6. CONCLUSION ON ROLE OF EPARTICIPATION IN STRENGTHENING CITIZENS POLITICIANS INTERACTION IN ENVIRONMENTAL ISSUES

Public involvement can improve the quality of decision-making, raise citizen awareness of environmental issues and increase public understanding of the decision-making process. Public involvement can be maximized by Internet-based approaches and the Web should be seen as a means of enhancing current practices. Geographic information and Internet technologies used for eParticipation may be employed as positive tools for increasing the transparency of decisions on environmental issues. Simultaneously, citizens will be provided with means to become familiarized and involved with the functioning of democratic decision making processes, to monitor the implementation of EU environmental law in their countries and to express their concerns over environmental matters affecting them. Thus, eParticipation can help to create more informed local, regional and national government decision-making processes by incorporating diverse opinions, values and ideas and by resulting in direct, immediate knowledge of the environmental conditions from the community and citizens. It also helps to defuse conflict or opposition to particular governmental actions, to build broad-based consensus for environmental programs and garner more support for their implementation. Of course, governments need to adapt their structures and processes to ensure that the results of online consultations are analyzed, disseminated and used.

ACKNOWLEDGEMENT

U@Marenostrum - Strengthening Public Participation for water protection and management is co-funded by the European Commission under the eParticipation Preparatory Action (EP-08-01-05)

REFERENCES

Aitken, S. Public participation, technological discourses, and the scale of GIS. In *Community Participation and Geographic Information Systems*, eds. W. Craig, T. Harris, and D. Weiner, pp. 357-366. London: Taylor & Francis, 2002.

Barndt, M. "Public participation GIS: Barriers to implementation." *Cartography and Geographic Information Systems* 25/2: 105-112, 1998.

Barndt, M. "A model for evaluating public participation GIS." In *Community Participation and Geographic Information Systems*, eds. W. Craig, T. Harris, and D. Weiner. London: Taylor & Francis. 346-356, 2002.

Blixt, P. *The eParticipation Initiative: Policy Context and Commission Actions – DG INFSO*, eParticipation Information Day August 31, Brussels, 2006.

Carver, S. Participation and geographic information: a position paper. Available at http://www.geog.leeds.ac.uk/papers/01-3/01-3.pdf , 2001.

Clark M J. GIS- Democracy or Delusion? *Environment and Planning A* 30(2): 303-316. *Information Systems* 28(3): 195-208, 1998.

Coleman, S. and Gøtze, J. *Bowling Together: Online Public Engagement in Policy Deliberation*, Hansard Society and BT, 2001

Coleman, S. and Norris, D. *Forum Discussion Paper No 4: A New Agenda for e-Democracy*, 2005.

Craig, W., Harris, T., and Weiner, D. (Eds.). *Community Participation in Geographic Information Systems.* London: Taylor & Francis, 2002.

Macintosh, A. 'Characterizing e-participation in policy-making', *Proceedings of the Thirty-Seventh Annual Hawaii International Conference on System Sciences, (HICSS-37)*, Big Island, Hawaii, 5–8 January, [online] http://csdl2.computer.org/comp/proceedings/hicss/2004/2056/05/205650117a.pdf, 2004a

Macintosh, A. and Whyte, A. 'Towards an evaluation framework for eParticipation', *Transforming Government: People*, *Process and Policy*, Vol. 2, No. 1, pp.16–30, 2008

Macintosh, A., Davenport, E., Malina, A. and Whyte, A. 'Technology to support participatory democracy', in Grönlund, Å. (Ed.): *Electronic Government: Design, Applications, and Management*, Idea Group Publishing, Umeå University, Sweden, 2001.

Meredith Th. Democracy and Participation in Environmental Decision-Making *Human Ecology Review, Vol. 9, No. 1, 2002*

Nyerges, T. and P. Jankowski. Toward a Participatory Geographic Information Science, in B. Wharf, K. Hansen, and D. Jannelle (eds.), *WorldMinds: 100 Geographic Problems*, Kluwer, Amsterdam, pp. 535-539, 2004

Nyerges, T. and P. Jankowski, Participatory Geographic Information Science, in Ari-Veikko Anttiroiko and Matti Malkia (eds.) *Encyclopedia of Digital Government*, Idea Group Inc, v. III, 1314-1318. , 2006.

Sieber, R., Public participation geographic information systems: A literature review and framework. *Annals of the Association of American Geographers* 96(3): 491-507, 2006.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pilmmann, I.Holoubek, R. Legat (eds.)
Masaryk University, 2009

A Transactional Environmental Support System for Europe

R. Kenward[a], **B. Manos**[b], **S. Arampatzis**[c], **J. Papathanasiou**[d]

[a] *NERC Centre for Ecology & Hydrology, Benson Lane, Crowmarsh Gifford, Wallingford OX10 8BB, UK (reke@ceh.ac.uk)*

[b] *Faculty of Agriculture, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece (manosb@agro.auth.ac.gr, jasonp@gen.auth.ac.gr)*

[c] *Tero Ltd, 21 Antoni Tristi St., Building B, 57001 Thessaloniki, Greece (info@tero.gr)*

[d] *Department of Marketing and Operations Management, University of Macedonia, Edessa, Greece (jasonp@uom.gr)*

Abstract: Human society developed through technological revolutions. Centralisation began with agriculture. An industrial revolution then led to strongly urbanised societies. Unfortunately, agriculture, industry and urbanisation now threaten the biosphere on which humans depend. Biodiversity conservation focussed on protection until Convention on Biological Diversity also emphasised sustainable use and an ecosystem approach, with local knowledge, monitoring and empowerment for adaptive management. However, conservation through use of biodiversity and ecosystem services is an aspect of multifunctional land-use that is socio-economically and ecologically complex. It requires knowledge support for local decisions that are made much more often than statutory environmental assessments, and thus too frequently for all to be guided by experts. The TESS project (www.tess-project.eu) is designing a system to collate and automate local delivery of all ways to leverage biodiversity enhancement throughout Europe, aiming to (i) predict impacts of small-scale actions on incomes and biodiversity, (ii) monitor results of the decisions that follow such prediction and (iii) inform central assessors to enable appropriate tuning of regulatory and fiscal incentives. If the information revolution can rebuild local knowledge in cooperation with central decision-makers, perhaps it can maintain a diversity of species, environments and cultures that sustains humanity.

Keywords: Biodiversity; Ecosystem Services; Monitoring; Adaptive Management; EIA.

1. CHANGING HUMAN IMPACTS AND CONSERVATION CONCEPTS

Human societies developed agriculture separately in all the major continents, leading to settlements, specialisation and varying degrees of centralisation [Bronowski 1973]. Further increase in density and urbanisation accompanied industrialisation, which started in Europe more than two centuries ago, spreading to North America a century later and recently to many more countries. A consequence of the specialisation and centralisation was rapid advance in scientific knowledge, so that pressures of human population growth have been accompanied for more than a century by increased understanding of links between human societies and their environment [Diamond 1997]. Concepts at first academic, about anthropogenic climate change [Arrhenius 1896], resource depletion [Ehrlich & Ehrlich 1970] and biosphere feedback mechanisms [Lovelock 1979], have become wide concerns.

In Europe, adverse development can be constrained (under 85/337/EEC) after Environmental Impact Assessment (EIA) at local level [e.g. Treweek 1999] and more recently (under 2001/42/EC) following Strategic Environmental Assessment (SEA) at higher level [Wood & Jeddow 1992]. These high-level directives, and protection of areas (e.g. Habitats Directive, 92/43/EEC), may halt loss of biodiversity by 2010 at continental level. However, the current system of assessment is bottlenecked by dependence on experts, which limits application to large or severe cases and can also prejudice repeatability in conflictive ways [Therivel 2004]. The complexity of assessment is daunting, especially when Sustainability Impact Assessment (SIA) includes social and economic sustainability factors as well as the environment [Jacobs & Sadler 1989, Therivel & Minas 2002]. Moreover, a challenge not met by the current system of assessments, is to influence the myriad decisions made by individuals at local level, on what to remove or plant and how and when to manage it. Decisions that are made for farm fields and gardens are small-scale individually, but summate to change the environment.

The limitations of protective measures were foreseen in the Convention on Biological Diversity (CBD) that came from the Rio de Janeiro Summit in 1992. CBD covers protection of species and habitats in just 2 of 19 substantive Articles, but considers use of the components of biodiversity in 13 of those Articles, notably recommending to: “protect and encourage customary use of biological resources in accordance with traditional cultural practices that are compatible with conservation or sustainable use requirements” (Article 10); and “adopt economically and socially sound measures that act as incentives for the conservation and sustainable use of components of biological diversity” (Article 11).

The complementation of protection by “incentive-based conservation” [Hutton & Leader-Williams 2003] is important, because globally the majority of land is not protected. Reserve creation peaked in the 1980s [Pretty 2002] and extension much past the current 12% of land protected globally may not be socio-economically feasible, despite estimates from species-area curves indicating that retention of biodiversity requires the application of conservation measures to some 50% of the land surface [Soulé & Sanjayan 1998]. The problem is that protection of habitats and species creates opportunity costs [Swanson 1992, Norton Griffiths 1995]. The resulting reduction in jobs or incomes can cause conflicts for farmers, foresters and fishers [e.g. Redpath et al. 2004] and even local poverty [Adams et al. 2004], with a resulting dichotomy of land use into areas (i) exploited intensively to produce food and other materials or (ii) protected for science, aesthetics or to prevent extinction of species. The alternative, a dual approach to conservation [Inamdar et al. 1999], envisions a “biodiversity friendly mosaic of land uses driven by the livelihoods that are derived from the sustainable use of wild living resources, instead of landscapes with small islands of biodiversity in a sea of agriculture” [Hutton & Leader Williams 2003].

Fortunately, Europe is moving towards a broader basis for conservation than a focus on protection of species and habitats. The Natura 2000 network specifically includes provision for use of wild resources, for example through hunting. Earlier fears that consumptive use of biodiversity risks a “tragedy of the commons” [Hardin 1968] are being replaced by community-based conservation [Berks et al. 1989, Ostrom et al. 1999]. Protection remains an important indication that society values wild species, and protected areas can play an important role for supporting core populations that render harvest more productive in surrounding areas [Roberts et al. 2002]. However, extractive use too is seen as desirable in many nature reserves, to maintain human practises that preserve habitats [Getz et al. 1999].

2. ECONOMICS FOR CONSERVATION THROUGH USE OF BIODIVERSITY

For a global stocktaking of human impacts on the biosphere, the Millennium Ecosystem Assessment (2004) introduced a further utilitarian concept: 4 types of ecosystem services. “Supporting” and “regulating” services benefit society as a whole and can therefore be considered public goods, to be sustained by public funding (e.g. agri-environment services in the second pillar of a revised Common Agricultural Policy), whereas “provisioning” and “cultural” services often benefit particular sectors in society and become private goods. However, supporting and regulating services do not necessarily require high biodiversity, partly because humans can fill the consumptive role of many other species (especially predators), and there is pressure to reduce funding for a CAP no longer giving cheap food.

Unfortunately "provisioning" through forestry and agriculture also tend to become so intensive that biodiversity suffers, while commercial or subsistence use of wild plants, fish and bush-meat becomes unsustainable. However, the "cultural" ecosystem services have a strong potential for conserving biodiversity. This is not simply a matter of eco-tourism, in which high-carbon travel may associate with high pressure on local water and other resources, but also of local communities gathering flowers, angling or hunting as much for recreation as for food, and developing rules to keep the service sustainable. For more than two millennia, wildlife reserves have been created for hunting [Gadgil & Guha 1992], with modern recreational hunters and anglers adding closed seasons, quotas, catch and release. Religions that preserve wildlife species are prevalent where human populations are dense, and the Q'ran records early sanctuaries at Makkah and Medina.

Where land is relatively unproductive, cultural use of wild resources frequently competes effectively with intensive uses, for example where hunting is more economic than livestock farming in southern Africa [papers in Prins et al. 2000], and hunting has restored endangered wildlife populations through management and reintroduction much more widely [papers in Dickson et al. 2009]. However, where soil fertility and climate combine to give high productivity, as in Europe, there is less land on which sustainable use of wild resources is more cost effective than intensive cultivation or other development. Moreover, the residual low-productivity areas tend to be refuges for rare species, which can (ironically) inhibit conservation through sustainable use of wild resources. Grouse moors are an example [Redpath et al. 2004]. In rapidly developing areas, fertile and accessible land which is not either protected, valuable for recreation or covered by construction tends to be used mainly for cultivation. New crops, such as bio-fuels, open more areas to cultivation and, as cultivation intensifies, biodiversity is lost.

There is a risk that polarised attitudes to polarised landscapes are again focusing conservation efforts on protected areas, while scope for restoring biodiversity elsewhere is overlooked. Studies repeatedly show that small de-intensification measures can have major impacts at little cost. Newton (2004) identified the main factors associated with decline of 30 bird species as: (i) weed control, (ii) early ploughing, (iii) grassland management, (iv) intensified stocking, (v) hedgerow loss & predation. All can be addressed in ways that produce fractional reductions in yield. An early example is a small reduction in cereal crop yields when headland-edges are left unsprayed, which increases abundance of game birds and other wild fauna and flora [Boatman & Sotherton 1988].

Then, if land gives annual income per hectare of I from intensive production, the reduced income from constrained use C should be acceptable if compensated by income U per hectare from use of wild resources:

$$C + U \geq I$$

This equation (Kenward & Garcia Cidad 2005) indicates how ecology can be combined with economics for decisions about land-use, and can also accommodate social aspects if measurement of C, U and I is extended from income per hectare to employment or quality of life measures. Moreover, just as the equation $C + U \geq I$ uses ecosystem-service value to leverage biodiversity through minor constraints on crop production, so can U be leveraged with stewardship subsidies S to maximise scope for conservation (i.e. $C+S+U \geq I$). There are already European examples where $C+U \geq I$ without subsidies. For example, where food shortage for deer in conifer plantations results in an uneconomic venison harvest and severe bark-stripping, a small loss of timber through including some deciduous trees can be more than offset by gain in value of deer and reduced damage [Reimoser & Reimoser 1997].

This leveraging approach can in principle be applied to all land, but is there enough value in use of wildlife resources for conserving much biodiversity? The most rigorous data come from surveys of spending on wildlife-associated recreation that are run at 5-year intervals by the United States Department of the Interior Fish and Wildlife Service and United States Department of Commerce Census Bureau. The latest national survey (USDI, FWS & USDC 2007) estimates that 88 million US adults (38% of adults) watched (71m), fished

(30m) and hunted (13m) wildlife in 2006, spending $122 billion. That represents $155 for each of the 774 million hectares of the USA.

In Europe, an FP6 project on Governance and Ecosystem Management for the Conservation of Biodiversity (GEMCONBIO) conducted a large case study on uses of wild biodiversity in the European Union, also during 2006 (www.gemconbio.eu). With help from the Federation of Associations for Hunting and Conservation of the EU, the European Anglers Alliance, the European Council for Conservation of Fungi, and Birdlife International, a high proportion of the 27 EU states in 2006 were covered by the survey for hunting (96%) and angling (64%), 81% for bird-watching and 42% for collecting fungi; other surveys increased the coverage to 100% for hunting and 94% for angling (Table 1).

Table 1. Participated and spend on wildlife-related activities in the EU.

	Proportion of EU population surveyed	Participants (grossed-up) millions	Annual spend € billions
Hunting	**96-100%**	**6.6**	**16**
Angling	**64-94%**	**23**	**19**
Collecting: Fungi	**42%**	[45][1]	
Plant Products	**7%**	[135][1]	
Bird-Watching	**81%**	**6.2**	{8}[2]

[1,2]unreliable due to [low survey%]{few spend data}

With participation in hunting and angling also estimated from licence data, estimates that about 7 million Europeans are recreational hunters and 23 million are anglers are probably reliable. With data from 10-14 countries on spending, it can be estimated that they spend about €35 billion annually, or at least €40 billion if (less reliable) estimates from bird-watching are included [Kenward & Sharp 2008], and that perhaps a quarter of the 490 million EU citizens gather fungi and plant products. This is equivalent to at least €121 for each of the 331 million hectares of the EU, equivalent to $181/ha at the exchange rate in late 2007. In the UK alone, a survey in 2002 estimated annual income from a wide range of wild resources (including collection of plant products and fungi but excluding released game) at €7.2 billion, which was 30-50% the value of UK agricultural production and accounted for some 58,000 jobs [IUCN-UK & ESUSG 2004].

Private recreational spending on wild resources encompasses equipment, accommodation and travel as well as use of land, but there is clearly scope for funding to benefit diversity of wild resources. It was also clear from the survey that although hunting and angling are well regulated and contributing to conservation across Europe, such that ungulate populations are stable or expanding in every surveyed state, much less attention has been paid to the potential contributions from those collecting fungi and plant products, while the relative economic contribution from wildlife-watching is much lower than in the US.

3. GOVERNANCE FOR CONSERVATION THROUGH USE OF BIODIVERSITY

In the same year as the Millennium Ecosystem Assessment was published, the parties to CBD committed at the 7th Conference of Parties to two documents, containing 12 "Malawi Principles" for the "Ecosystem Approach" (CBD VII/11) and 14 "Addis Ababa Principles and Guidelines (AAPG)" on sustainable use of biodiversity (CBD VII/12). These principles give as much consideration to social issues and economics as they do to ecological issues, as also recommended at the World Summit on Sustainable Development of 2002.

A recent review of axioms and principles for sustainable use concludes that benefit-sharing, adaptive management, local empowerment and ecosystem maintenance are almost universal, with integration, multiple scales and realising values common to most [Wall & Kernohan in press]. Frequent recommendation of these principles reflects much qualitative analysis of governance. In particular, adaptive management from local knowledge and monitoring has long been recommended for conservation [Holling 1978]. Demonstration of quantitative associations between these principles and ecosystem services is rare. However, the GEMCONBIO project used 36 cases, from local to international level, to show highly significant positive associations of ecosystem service sustainability with adaptive management. The project also found positive linkage of ecosystem services and biodiversity with external knowledge leadership and community-based management, with regulations playing positive and negative roles.

In 2007, parties to the 1979 (Bern) Convention on the Conservation of European Wildlife and Natural Habitats approved a charter a based on 12 simple recommendations condensed from the Ecosystem Approach and AAPG commitments of CBD (TVPS(2007)7). A charter is a document that agrees responsibility of government towards citizens (effectively conferring rights) as well as responsibility of citizens. Therefore the charter not only has guidelines for users of wild resources, for instance on monitoring and adaptive management, but also for regulators at all levels so that they too can encourage conservation through use of biodiversity. The charter is on hunting and biodiversity, but the 12 recommendations do not specify hunting and are therefore equally applicable to all use of wild species, whether through angling, gathering wild flowers or collecting fungi.

Although the recommendations in the charter are simpler and fewer than in the documents on which they are based, each recommendation involves many guidelines. Moreover, in practical application of the economic considerations in part 2, many crops can be grown on land in different ways, many types of de-intensification applied for different species, potentially using funds from public sources and many private activities applied in different ways. Compared to the simplicity of protective regulation (i.e. to not use this land or not to take that species), conservation through use of biodiversity is complex indeed.

4. TESS: A NEW TOOL FOR COMPLEX SOCIO-ECOLOGY

Much of the damage to global ecosystem services has occurred in the last 50 years. This reflects human population increase, and also improved technology for felling trees and tilling land, but also governance measures. Thus, 50 years of subsidies at continental and state level have successfully driven production based on very few crop species in Europe, and hence on minimal biodiversity. Commercially driven homogenisation of diverse local land-use continues to degrade ecosystem services that sustained Europeans for centuries [Pretty 2002]. Species whose dynamics and colonisation operates at small scale have disappeared through habitat loss and fragmentation, so that biodiversity has declined drastically at local level [e.g. Paine & Pienkowski 1997, Thomas et al. 2004].

Yet, over the same 50 years, scientific knowledge of the environment has also advanced rapidly, so that “Paradoxically we are not limited by lack of knowledge but failure to synthesise and distribute what we know” [Pimm et al. 2001]. As noted above, the synthesis must handle many types of species, land use, wild resource use and regulation based on the need for sustainability, so the calculation process becomes complex. Sophisticated decision support becomes essential. However, during the last 30 years, human ability to integrate information, handle complex calculation and deliver predictions as decision support has increased dramatically, through the use of computers. Forecasts from empirical models, which neglect causation, are being replaced by more accurate prediction from individual-based models that incorporate behavioural processes [Goss-Custard & Sutherland 1997]. Models can be spatially specific through linkage to habitat and socio-economic data as cells in geographic information systems (GIS).

The internet is the way to collate the extensive ecological knowledge that is currently fragmented across Europe, as recognized by important data collation and standardisation initiatives such as GMES, INSPIRE and SEIS. Information from geo-referenced databases at national level is already used by government experts for SEA, and by private consultancies to meet EIA requirements of government in large development projects. Thus the internet increasingly delivers knowledge, computed by experts, to implement commitments of CBD parties in 2004 on environmental assessment. The need now is to include socio-economic data, to meet other commitments at CBD/VII, on incentives, sustainable use and the ecosystem approach. Inclusion of socio-economic data is an asset rather than a complication, because private spending on wildlife-related activities is large and often requires or creates high biodiversity. Such private funding needs to be tapped and used cost-effectively for conservation of biodiversity in Europe.

However, the task of applying private recreational payments to enhance biodiversity, for example through minor constraints on crop production, needs to be done at local level. So do other projects that can offset wider commercial pressures (e.g. farm shops), and tuning to local conditions of other small de-intensification measures (e.g. headlands), public works (e.g. road verges) and gardens that can benefit biodiversity at minimal cost. Moreover, it is the myriad of decisions made locally by individuals (on what and when to plant or remove, what to consume or discard or how to travel) that gradually change the biosphere.

There are can never be enough human experts to handle all the complex knowledge needed at local level to handle decisions that fall outside the scope of SEA and EIA. But what if those individual decisions too could be guided automatically via the internet, in ways as subtle as the red and green underlining in a word processor? An electronic farm-plan could predict income from crops and flash for a buffer strip to reduce nitrate run-off (Table 2). An architect’s plan could colour the best sites for solar cladding, or a garden plan offer rewards for a carbon-sequestration scheme. A beep on the GPS-enabled tractor could warn where to avoid mowing rare plants or bird’s nests.

Table 2. Examples in modes of operation of an environmental decision support system.

Scale	Question	Operations
Field	Is it too early for the Nymphalis butterfly larvae to cut these nettles in our amenity area now?	*Map on hand-held remote communication device with GPS-auto-location capability.*
Farm	If I use my land like this in future, what happens to my income, game bags and nitrate run-offs?	*Completion of electronic farm plan attracts colour coding and comment on proposed mitigations.*
Parish	How do we route this path to optimise views while minimising erosion and wildlife disturbance?	*GIS-based modelling with 2D/3D views on desk-top PC in local community centre.*
Region	If trends in land-use continue for 20 years, how will BAP targets be affected? Can subsidies ameliorate?	*Statistics plot in government department after country wide distributed parallel processing.*

What if the results of all those decisions could also be reported to central planners, as a GIS for species and habitats in ever-increasing detail, to enable more sophisticated SEA and SIA in exchange for the decision support that benefits local livelihoods and biodiversity? At local level, baseline monitoring and continuing assessment over wide areas could also solve several problems with EIA, enable ‘pay by results’ to replace ‘pay for process’ subsidies [Ferrano & Kiss 2002] and stimulate interest in widespread biodiversity restoration (thereby perhaps reducing need for high-carbon visits to distant reserves).

The application of local knowledge for adaptive management inevitably faces the challenge of building a functioning link between local communities and central decision-makers. With the co-financing of the European Commission, an FP7 project is designing a Transactional Environmental Support System (TESS). The aim is to find how best the important data collation and standardisation initiatives such as GMES, INSPIRE and SEIS can be combined with the abundant but disparate environmental research across Europe, for predictive modelling to inform not only SEA and EIA, but also the myriad individual decisions. This will enable the regional authorities and local stakeholders to simulate different scenarios and design policies for optimal decisions.

TESS contends that local communities can restore environments if they are enlightened, empowered and aided by policy-makers and society as a whole, with the use of ICT services. The reasoning is that:

- central planners can collate complex knowledge and incentives to assist local decisions;
  - but they need local information to monitor and adapt their knowledge and incentives policy;
- local managers must gather local information to make and monitor their decisions;
  - so they can exchange this local data for the complex knowledge that benefits their livelihoods;
- the huge volume of local-centre exchanges will need an automated support system.

The automated support system will handle environmental data, but also include information on market economics and government incentives in a way that is both user-friendly and socially integrated. Social integration can be planned partly by survey and partly by trials in local communities across Europe. TESS trials will use GPS-enabled PDAs to map species and habitat, and local projects to benefit biodiversity and livelihoods. Other issues include standardisation of data and models, data security, scaling, and scope for e-commerce.

5. CONCLUSIONS

1. Conservation through use of biodiversity requires complex local decisions, based on species requirements and bio-socio-economic considerations for multi-functional land use.
2. The internet provides a way to bring together widely dispersed predictive models, and to collate complex environmental data, for providing knowledge leadership to local level.
3. The knowledge needs to be delivered to land-managers, in the form of context-adaptive support for decisions on livelihoods and biodiversity, as much as to government planners.
4. Local decision support requires local records of species and habitats, which central government also needs mapped in order to manage the decision support system adaptively.
5. A system in which central decisions makers can exchange decision support to local level for local knowledge would facilitate statutory environmental assessment as well as myriad small local-level decisions that summate to change the human environment.

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n°212304

REFERENCES

Adams, W.M. et al., Biodiversity conservation and the eradication of poverty, *Science 306*, 1146-1149, 2004.

Arrhenius, S., On the Influence of Carbonic Acid in the Air upon the Temperature of the Ground, *Philosophical Magazine 41*, 237-276, 1896.

Berkes, F., Feeny, D. McCay, B. J., Acheson, J. M. The benefits of the commons, *Science 340*, 91–93, 1989.

Boatman, N.D. and N.W. Sotherton, Agronomic consequences and costs of managing field margins for game and wildlife conservation, *Annals of Applied Biology 17*, 47-56, 1988.

Bronowski, J., *The ascent of man*. BBC Publications, London, UK, 1973.

Diamond, D., *Guns and germs and steel: the fates of human societies*, Jonathan Cape, London, UK, 1997.

Dickson, B., J. Hutton, J. and W.M. Adams (eds), *Recreational hunting, conservation and rural livelihoods*, Wiley-Blackwell, Oxford, UK, 2009.

Ehrlich, P.R. and A.H. Ehrlich. *Population, resources, environment*. Freeman, San Francisco, 1970.

Ferrano, P.J. and A. Kiss, Direct payments to conserve biodiversity, *Science 298*, 1718-19, 2002.

Gadgil, M. and R. Guha, *This Fissured Land: An Ecological History of India*, Oxford University Press, Delhi, India, 1992.

Getz, W.M. et al., Sustaining natural and human capital: villagers and scientists, *Science 283*, 1855-1856. 1999.

Goss-Custard, J.D. and W.J. Sutherland, Individual behaviour, populations and conservation. In: J.R. Krebs and N.B. Davies (eds), *Behavioural Ecology: An Evolutionary Approach*, Blackwell Science, Oxford, UK, pp. 373-395, 1997.

Hardin, G., The tragedy of the commons, *Science 168*, 1243-8, 1968.

Holling, C.S., *Adaptive environment assessment and management*, Wiley, New York, 1978.

Hutton, J.M. and N. Leader-Williams, Sustainable use and incentive-driven conservation: realigning human and conservation interests, *Oryx 37*, 215-226, 2003.

Inamdar, A., H. De Jode, K. Lindsay and S. Cobb, Capitalising on nature: protected area management, *Science 283*, 1856-1857, 1999.

IUCN-UK & ESUSG. *Wealth from the Wild: a Review of the use of wild living resources in the United Kingdom*. UK Committee of the IUCN (the World Conservation Union) and European Sustainable Use Specialist Group of IUCN/SSC, 2004.

Jacobs, P. and B. Sadler, (eds), *Sustainable development and environmental assessment: perspectives on planning for a common future*, Canadian Environment Assessment Research Council, Quebec, Canada, 1989.

Kenward, R.E. and V. Garcia Cidad, Innovative approaches to sustainable use of biodiversity and landscape in the farmed countryside, pp 565-589 in *UNEP High-Level Pan-European Conference on Agriculture and Biodiversity*, Council of Europe, Strasbourg, France, 2005.

Kenward, R. and R. Sharp, Use Nationally of Wildlife Resources across Europe (UNWIRE), pp. 82-86 in Manos, B. and J. Papathanasiou (eds) *GEMCONBIO: Governance and Ecosystem Management for Conservation of Biodiversity*, Aristotle University of Thessaloniki, Greece, 2008.

Lovelock, J.E., *Gaia: a new look at life on earth*, Oxford University Press, UK,1979.

Newton, I. The recent declines of farmland bird populations in Britain: an appraisal of causal factors and conservation actions. *Ibis 146*, 579-600, 2004.

Norton-Griffiths, M. and C. Southey, The opportunity costs of biodiversity conservation in Kenya, *Ecological Economics 12*, 125-139, 1995.

Ostrom,E., J. Burger, C.B. Field, R. Norgaard and D. Policansky, Revisiting the commons: local lessons, global challenges, *Science 284*, 278-282, 1999.

Pain, D.J. and M.W. Pienkowski (eds) *Farming and birds in Europe*, Academic Press, London, UK, 1997.

Pimm, S.L. et al., Can we defy nature's end? *Science 293*, 2207-2208, 2001.

Pretty, J., *Agri-culture- reconnecting people, land and nature*, Earthscan, Cambridge, 2002.

Prins, H.H.T., J.G. Grootenhuis and T.T Doan (eds), *Wildlife conservation by sustainable use*, Dordrecht, Netherlands, 2000.

Redpath S.M., B.E. Arroyo, F.M. Leckie, P. Bacon, N. Bayfield, R.J. Gutiérrez and S.J. Thirgood, Using Decision Modeling with Stakeholders to Reduce Human–Wildlife Conflict: a Raptor–Grouse Case Study, *Conservation Biology 18*, 350-359, 2004.

Reimoser, F and S. Reimoser, *Zeitschrift fur Jadgwissenschaft 43*, 186-196, 1997.

Roberts, C.M., J.A. Bohnsack, F. Gell, J.P. Hawkins and R, Goodridge, Marine reserves amd fisheries management, *Science 295*, 1234-1235, 2002.

Swanson, T., *The international regulation of extinction*, Macmillan, London, UK, 1992.

Soulé, M.E. and M.A. Sanjayan, Conservation targets: do they help? *Science 279*, 2060-2061, 1998.

Therivel. R., *Strategic Environmental Assessment in Action*, Earthscan, London, UK. 2004.

Therivel, R. and P. Minas, Ensuring effective sustainability appraisal, *Impact Assessment and Project Appraisal 20*, 81-91, 2002.

Thomas, J.A. et al, Comparative Losses of British Butterflies, Birds, and Plants and the Global Extinction Crisis, *Science 303*, 1879-1881, 2004.

Treweek, J., *Ecological Impact Assessment*, Blackwell, Oxford, UK,1999.

USDI, FWS & USDC, *2006 National survey of fishing, hunting and wildlife-associated recreation*, US Department of the Interior Fish and Wildlife Service, Fish and Wildlife Service and US Department of Commerce Census Bureau, Washington DC., 2007.

Wall, W.A. and B.J. Kernohan, Key components of conservation-hunting programs and their relationship to wildlife populations, ecosystems and people, in *Proceedings of 3rd International Wildlife Management Congress*, Christchurch, New Zealand, in press.

Wood, C. and M. Jeddow, Strategic Environmental Assessment: EA of Policies, Plans and Programmes, *Impact Assessment Bulletin 10*, 3-21, 1992.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

The GIGAS project – an action in support to GEOSS, INSPIRE, and GMES

Eva Klien[a], Alessandro Annoni[b], Pier Giorgio Marchetti[c]

[a] *Fraunhofer Institute for Computer Graphics Research (IGD), Fraunhoferstr. 5, 64283 Darmstadt, Germany (eva.klien@igd.fraunhofer.de)*

[b] *Joint Research Centre, Institute for Environment and Sustainability, Spatial Data Infrastructures Unit, 21027 Ispra, Italy (alessandro.annoni@jrc.it)*

[c] *European Space Agency, Via Galileo Galilei, 00044 Frascati, Italy (pier.giorgio.marchetti@esa.int)*

Abstract: The three initiatives GEOSS, INSPIRE and GMES share commonalities, including their focus on environmental policy support, use of geomatic and geographic information, their Europe- or worldwide dimension, their reliance on international standards and the advanced Spatial Data Infrastructures that are needed for their implementation. At the same time however, each initiative follows its own timeline and approach for technical development, thereby risking evolving into separate, incompatible services and not profiting from the benefits of a common approach. The GIGAS project was launched with the specific aim to assess and address interoperability gaps and opportunities for establishing bridges between the initiatives. For this, GIGAS makes use of a formal and structured approach that is driven by consensus. To support the consensus building process effectively, one of the core objectives of GIGAS is the set up of a representative and sustainable stakeholder forum as central communication platform.

Keywords: European Spatial Data Infrastructure, GEOSS, INSPIRE, GMES

1 INTRODUCTION

Timely, reliable and relevant information on the state of the environment is essential for sound policies. The awareness about the need for quality geo-referenced information to support understanding of the complexity and interactions between human activities and environment is growing at national and European level. This includes information on how the climate is changing, whether European waters are improving and how nature is reacting to pollution and changing land use. Such information should be made available to all and be easily understood [European Communities 2008].

In this context, the objective of the Shared Environmental Information System (SEIS) is to tie better together all existing data gathering and information flows using modern tools such as the internet and satellite technology. The objective is also to move away from paper based reporting and reports to a system where data is made available to the policy makers and users in an open and transparent way.

The three existing initiatives GEOSS (Global Earth Observation System of Systems), INSPIRE (Infrastructure for Spatial Information in Europe) and GMES (Global Monitoring for Environment and Security) are of very different nature but share commonalities, including their focus on environmental policy support, use of geomatic and geographic information, their Europe- or worldwide dimension, their reliance on international standards and the advanced Spatial Data Infrastructures that are needed for their implementation. In this respect, all three initiatives are significantly contributing to the implementation of SEIS as a basis for improving sharing of environmental-related data and information and the provision of services. However, at the same time each initiative follows its own timeline and approach for technical development, thereby risking evolving into separate, incompatible services and not profiting from the benefits of a common approach.

In mid 2008, the GIGAS project (GEOSS, INSPIRE, and GMES, an Action in Support) was launched with the specific aim to assess and address interoperability gaps and opportunities for establishing bridges between the initiatives. The consortium responsible for the project consists of technical representatives of the three initiatives (GEOSS, INSPIRE and GMES), representatives of three main standardisation bodies (ISO, OGC and CEN) and representatives of SME stakeholders and research institutes. The project is funded by the European Union for the duration of two years.

GIGAS makes use of a formal and structured approach to identify and analyse commonalities and interoperability gaps in architectures, standards and governance issues in the three initiatives. This study and the subsequent comparative analysis lead to a set of draft recommendations directed towards the initiatives and the main standardization bodies (OGC, ISO, CEN).

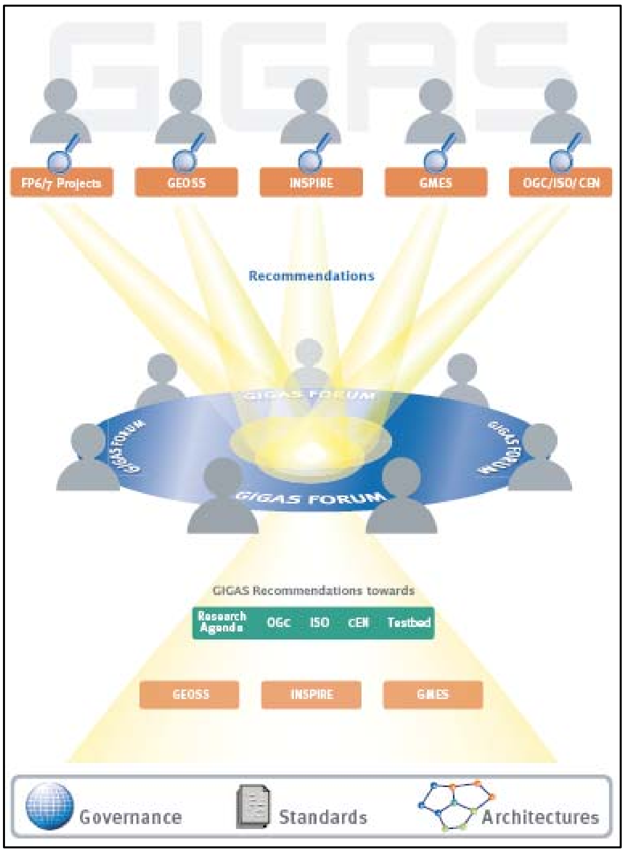


Figure 1. The GIGAS project.

Before any activity is derived from these draft recommendations, they will be revised through a broad consultation to achieve as wide a consensus as possible. The impact of GIGAS will largely depend on how successful we are in involving the initiatives' stakeholders into this consultation process. The core challenge thus relies in setting up a representative stakeholder forum to enable interested parties to assess, comment and support the recommendations that are proposed by the GIGAS Consortium.

To ensure that the recommendations are sufficiently discussed and agreed upon by stakeholders and that the recommendations reflect all their concerns, a consensus building procedure is defined and implemented with the help of two instruments:

- Regular Stakeholder Workshops for presenting and discussing the draft recommendations.
- The GIGAS Forum as communication platform for continuous discussion on any topic of interest in the context of harmonising European SDI initiatives. This online forum provides tools for consensus building to achieve broad acceptance on the GIGAS rec-

ommendations. This forum will be built up in the course of the GIGAS project and will persist as communication platform after the project has ended in June 2010.

The central position of the GIGAS Forum in the project is also reflected in Figure 1 that illustrates the general concept of GIGAS. The relevant stakeholders in this consensus building process are people who are involved in convergence and harmonization work in the context of GEOSS, INSPIRE, and GMES as well as people who are in the positions where recommendations are to be adopted. This includes decision makers in the three initiatives, standardization organisations, and research programs.

The rest of the paper is structured as follows. Section 2 introduces basic facts for GEOSS, INSPIRE and GMES and motivates the need for a project like GIGAS. Section 3 describes the consensus-driven approach that is performed in the GIGAS project. Section 4 provides more details on the central topic of the GIGAS Forum. In section 5, the current status of the project is reported and section 6 concludes with some final remarks.

2 TOWARDS HARMONISATION OF SDI INITIATIVES

In the following, we provide a short overview on the main objectives of the three initiatives GEOSS, INSPIRE and GMES. This overview is followed by a brief analysis of their commonalities and also of their differences – both on rather abstract level. The presented current situation motivates the need for a collaborative effort for convergence as proposed by GIGAS.

2.1 GEOSS (Global Earth Observation System of Systems)

The Group on Earth Observations is coordinating efforts to build a Global Earth Observation System of Systems (GEOSS) (see http://www.earthobservations.org/*).* GEO that is a voluntary partnership of governments and international organisations is constructing GEOSS on the basis of a 10-year Implementation Plan that runs from 2005 to 2015. The Plan define a vision statement for GEOSS, its purpose and scope, expected benefits, nine "Societal Benefit Areas" (disasters, health, energy, climate, water, weather, ecosystems, agriculture and biodiversity), technical and capacity building priorities and GEO governance. The vision for GEOSS is to realize a future wherein decisions and actions for the benefit of humankind are informed by coordinated, comprehensive and sustained Earth observations and information." To achieve this vision and stimulate broad use of GEOSS, the system of systems must provide ready access and improved interoperability for existing and future observation systems. The tools for access and interoperability are the core elements of the "Common Infrastructure" of GEOSS (GCI). These include ways through the web to identify and access services available. It also includes a list of interoperability agreements and best practices that ultimately allow users to employ global data sets without having to translate data from different national and regional systems. This common infrastructure includes the core components and functions that link the various resources of GEOSS together. The GCI consists of web-based portals, a clearinghouse for searching data, information and services, and registries containing information about GEOSS components, standards, and best practices.

2.2 INSPIRE (establishing an Infrastructure for Spatial Information in the European Community)

INSPIRE is a European Directive that lays down general rules to establish an Infrastructure for Spatial Information in Europe. It has been designed primarily for the purposes of Community environmental policies and policies or activities which may have an impact on the environment. INSPIRE will be built upon infrastructures for spatial information established and operated by the Member States so leaving to them a higher degree of freedom in designing their own architectures. The Directive must be transposed into national legislation by 15 May 2009. In some countries INSPIRE will be embedded in government acts whereas in other countries the infrastructure will primarily target the Environmental community. INSPIRE components include Metadata, Interoperability of Spatial Data and Spatial Data Services, Network services (discovery, view, download, invoke), Data and Service sharing and Coordination and measures for Monitoring & Reporting. The detailed technical provisions for the issues above will be laid down in Implementing Rules (IR), which are being developed in close collaboration with Member States' organisations and other stakeholders. The INSPIRE Commission Regulation on Metadata (EC) No 1205/2008 has been recently approved. The European Commission will operate the INSPIRE Geo-portal - an Internet-based facility that will provide access to Member States services. These services should allow the users to identify and access spatial or geographical information from a wide range of sources, from the local level to the global level, in an interoperable way for a variety of uses.

2.3 GMES (Global Monitoring for Environment and Security)

GMES (see http://www.gmes.info) is a common initiative of the European Commission (EC) and European Space Agency (ESA). initiated in 1998 and endorsed in 2001. The programme aims at ensuring European independence in critical data sources and is considered a European flagship programme after the Galileo navigation system. A challenge for GMES is to gather relevant spatial and in-situ data and to provide innovative, cost-effective, sustainable and user-friendly services. They will enable decision-makers to better anticipate or integrate crisis situation issues related to the management of environment and security. GMES can assist through improved prediction, monitoring and assessment capabilities, in the preparation of strategies to cope with natural hazards and human made disasters, thus contributing to the reduction of the resulting economic losses.

2.4 Commonalities and Discrepancies

The above descriptions of the three initiatives GEOSS, INSPIRE and GMES show an explicit overlapping between their objectives, missions and tasks.

Concepts as "interoperability" and "harmonisation" are the core elements of the different systems, which provide complementary approaches to similar goals and objectives. However, there are important differences in related architectures and in the internal emphasis given to adoption of common standards. In addition, the different processes with which the three programmes are run, in term of way of execution, actors involved and different time-frames, bring in the risk of a substantial divergence of the process outcome with duplication of efforts, delays in implementation time-lines, increases of costs, additional barriers in data access and exploitation.

In the scope of the three initiatives, the benefits for standardisation and interoperability can be summarized as follows:

- Manage and reduce cost of the initiatives' systems and operations
- Manage and reduce technical risks in the initiatives' systems and operations
- Ensure that technology drivers for the across initiative interaction are lead by agreed requirements
- Allow interoperability within and across organisations

With the exploitation of synergies among INSPIRE, GEOSS and GMES, the GIGAS project aims at enhancing the interoperability level between the three systems and to provide a stronger support to standardisation activities with a wider scope.

3 THE GIGAS APPROACH

The goal of the GIGAS project is to promote the coherent and interoperable development of the GMES, INSPIRE and GEOSS initiatives through their concerted adoption of standards, protocols, and open architectures. The main objectives include:

- Analyse the gaps between the different initiatives and propose strategies to overcome them
- Highlight best-practice examples from finished FP6 or ongoing FP7 projects relevant for the identified gaps
- Initiate a consensus process on a broad basis for public consultation and consensus building
- Shape the initiatives by providing short term action items
- Influence the relevant standardisation bodies to ensure the long-term action
- Provide an agenda for further strategic research areas to ensure investigation on the problems that are unsolved today

In the following we describe the consensus-driven approach in more detail that is employed in GIGAS. The approach is based on four basic steps: monitoring, comparison, discussion and consensus, and influence and shape, illustrated in Figure 2. This process will be iterated twice in the course of the project. Experiences from the first iteration will help to refine the scope and targets for the second iteration and to improve the overall methodology.

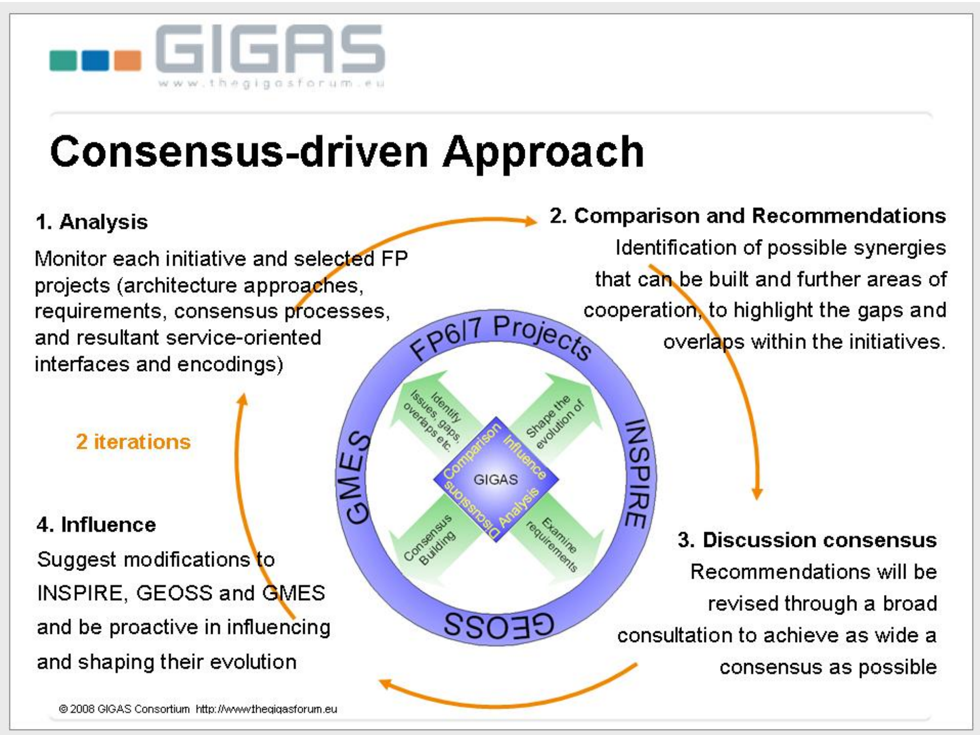


Figure 2: The consensus-driven approach in GIGAS.

Monitoring: This analysis task represents a learning phase to highlight the main issues. In fact to be able to influence and ensure coherent development of architectures, governance and standards, an understanding and deeper insight into the current architectures, governance and standards within the initiative is imperative. This ensures that issues, practices, protocols, tools and methodologies etc. that are currently in use are identified, analyzed and documented. This is one of the main tasks of GIGAS. The output of this process will be the commonalities, discrepancies, redundancies, or even inadequacies between the three initiatives and relevant FP6/FP7 projects.

Comparison: A full convergence will not be possible in the time frame to the project. For this reason more importantly is the identification of priority areas and possible synergies. A crucial task for GIGAS when examining the initiative is a comparison of the requirements and the adopted solutions to highlight the gaps and overlaps within the initiatives and areas where further cooperation can be achieved.

Discussion and consensus: Comparative analysis will lead to a multitude of issues that needs to be addressed in the initiatives as they stand currently. GIGAS seeks to initiate discussion on emerging priority issues in terms of governance, standards and architectural designs in order to reach common agreement. The highlighted issues are to be discussed and agreed upon using versatile consensus methodology and tools. This consensus-building effort is imperative, especially when standards are also involved.

Influence and shape: The real value of the GIGAS support action lies in the outcomes being used within the initiatives and relevant FP projects. Thus, GIGAS aim is to define a clear direction for further development of the three initiatives' architectures and standards. GIGAS intend to influence directly or indirectly the development of standards, architecture development and governance within the initiatives by targeting its consensus process outputs to the initiatives.

The consensus process that takes place in the third step of the approach is crucial for the overall success of the exercise. Therefore, a core challenge in GIGAS is to set up a representative stakeholder forum. Since the forum concept plays a major role, we will discuss the idea and purpose of the forum in more detail in the following section.

4 THE GIGAS FORUM

First, a sound and suitable procedure for consensus building has to be defined on the requirements that are delivered by the GIGAS Consortium. This includes the definition of clear and transparent rules for the procedures as well as adequate tools to support and monitor the process. Another important part of the definition is a categorisation of stakeholders to ensure that all relevant parties are sufficiently addressed and involved in the process [Craglia, 2008].

Second, to ensure that the recommendations are sufficiently discussed and agreed upon by stakeholders and that the recommendations reflect all their concerns, the consensus building procedure has to be implemented. Through the GIGAS Forum, the stakeholders are offered the possibility to express their needs, reservations and to influence the recommendation process.

The GIGAS Forum is designed as a communication platform for continuous discussion on any topic of interest in the context of harmonising European SDI initiatives. However, the first version launched will focus on tools for supporting the consensus building process as part of the GIGAS approach. Further, the stakeholders will have easy access to all required information in order to make a qualified decision. The relevant stakeholders in this consensus building process are people who are involved in convergence and harmonization work in the context of GEOSS, INSPIRE, and GMES as well as people who are in the positions where recommendations are to be adopted. This includes decision makers in the three initiatives, standardization organisations, and research programs. In summary, the main goals of the online forum are:

- To facilitate communication amongst the registered members to achieve a broad consensus on the recommendations coming from the GIGAS Consortium.
- To ensure that the recommendations are sufficiently discussed and agreed upon by stakeholders representing various interests, ensuring that GIGAS recommendations reflect all stakeholders' concerns.
- To facilitate communication and coordination with other relevant fora.

This forum will be built up in the course of the GIGAS project and will persist as communication platform after the project has ended in June 2010. If successful, this platform will be the key for connecting the European community and building a European voice.

5 PROJECT STATUS AND REQUEST FOR STAKEHOLDER INVOLVEMENT

The current output of the project refers to the first two steps of the GIGAS approach (Figure 2). This includes:

- A methodology for requirement analysis that has provided a formal and structured procedure for monitoring and comparing various topics of interests from different perspectives across the three initiatives [Marchetti et al., 2008].
- An analysis of scope and targets that influenced the selection of topics monitored by GIGAS partners in the first iteration [Annoni et al., 2008]
- Seven technical reports on the results of the monitoring activity on the following topics: Architecture interoperability, Catalogue, metadata and resource discovery, Sensor Planning Service, User Management, Ordering, Web Map Service, Observation and measurements.
- A Comparative analysis report that identifies the interoperability gaps across the initiatives for each topic and formulates draft recommendations on how to address them.

In the next phase, the draft recommendations have to be revised in a broad consultation process to achieve as wide a consensus as possible. From these recommendations then follow actions to shape the direction of the initiatives and to define a roadmap for future development, including the key research topics to be addressed to sustain the convergence of the initiatives. This phase that subsumes step 3 and 4 of the GIGAS approach (Figure 2) has been kicked-off with a Stakeholder Workshop on $28^{th}/29^{th}$ of January 2009 in Brussels, and it will be continued as open discussion in the GIGAS Forum. The necessary tools and procedures are provided on the GIGAS website: http://www.thegigasforum.eu.

6 CONCLUSION

GIGAS promotes the coherent and interoperable development of the GMES, INSPIRE and GEOSS initiatives through their concerted adoption of standards, protocols, and open architectures. The GIGAS Forum will facilitate discussion among the key stakeholder of the three initiatives, as well as other related programmes like the Shared Environment Information System for Europe (SEIS). The consensus-driven approach addresses key community stakeholders who are interested in strategies for influencing and shaping the evolution of European SDI initiatives (European focus) towards convergence.

GIGAS will contribute to the emergence of a collaborative information space for accessing and sharing distributed environmental resources in Europe. This will represent a milestone towards building a Single Information Space in Europe for the Environment

ACKNOWLEDGEMENTS

The work presented in this paper is the joint effort of the GIGAS consortium; the authors listed above are the members of the Executive Board of the project. The description of the entire consortium can be found at the project website:

http://www.thegigasforum.eu/partners.html.

The GIGAS project is a support action financed under the 7^{th} Framework Programme of the European Union (FP7 224274).

REFERENCES

Annoni, A, Fierens, F, and Lutz, M., D2.1 Requirement analysis methodology (part 2): Refine Scope and Target, GIGAS Deliverable, 2008.

Craglia, M, D3.1 Consensus Building Methodology, GIGAS Deliverable, 2008.

European Communities, Commission proposes to improve and streamline the European system for collecting, analysing and reporting environmental information. Press release, Reference: IP/08/185, Date: 05/02/2008, accessible at: http://europa.eu/rapid/pressReleasesAction.do?reference=IP/08/185, 2008.

Marchetti, P-M, Biancalana, A, Coene, Y. and Usländer, T, D2.1 Requirement analysis methodology (part 1): Define Methodology, GIGAS Deliverable, 2008.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Localised environmental and health information services (lenvis): a generic Decision Support Network

S. Loos[a], **A.H. Lobbrecht**[a]
[a] *HydroLogic BV, P.O. box 2177, 3800 CD Amersfoort, The Netherlands*
www.hydrologic.com

Abstract: There is a growing demand for real time and integrated environmental and health risk information. Provision of location-based services linked to the state of the environment at particular geographical locations is necessary for improving the quality of life. This is essential for mitigation of environmental-related health threats associated to water quantity and quality, and outdoor air pollution. The main goal of the lenvis project is to develop an innovative collaborative decision support network for exchange of location-based environmental and health services between stakeholders. Lenvis includes health indicators as integral part of the environmental management. Lenvis furthermore aims at supporting governments in communication with the general public and in particular to address the young generation Y, by entering popular digital information sources and social networks. This is done by development of a generic ICT solution on the basis of service-oriented-architecture (SOA), providing data, information and modelling services. The approach will be validated through test cases on surface water and outdoor air quality in the Netherlands, Portugal and Italy. Lenvis facilitates collaboration between different stakeholders, such as environmental protection agencies, health institutions and service providers, policy makers, citizens in general and environmental communities in Europe.
This paper presents the first ideas of the lenvis project which started in September 2008, specific technologies presented are subject to change during the project.

Keywords: Environmental information; Service Oriented Architecture; SOA; collaboration; water quantity; water quality; air quality; health indicator; monitoring; modelling; decision support.

1. INTRODUCTION

Human health and environmental quality are closely linked. Environmental pollution and in general poor environmental quality have well-established effects on human health and the quality of life. This increases the demand for real time and integrated environmental and health risk information. The lenvis project will integrate environmental and health information in the steps of identification and assessment of impacts, decision making and recommendations, having atmospheric pollution and water pollution as the risk factors for human health and creating a novel decision support network, capable to support stakeholders and citizens in their activities.

1.1 The lenvis goal

The overall goal of lenvis is to develop and validate an innovative collaborative decision support network (DSN) for exchange of location-based environmental and health services between stakeholders, for enhanced capacity to assess population exposure and health risks.

This network will specifically target European citizens and in particular the upcoming young generation-Y.
The lenvis project aims to fill the existing gap between environmental management and health management systems. This will be done by developing a generic ICT solution that combines service-oriented architecture (SOA) and user-centric approach by fusion of location-based environmental and health data, information and modelling services. This novel collaborative network is validated through test cases on fresh surface water and outdoor air quality in the Netherlands, Portugal and Italy. Lenvis also aims at bridging the socio-technology gap between the authorities / policy makers and the general public. This is achieved by use of communication channels appropriate to the upcoming generation-Y, using converged communications over IP as the basis of collaborative services and broadcasting capabilities. This young generation is branded as strongly decentralised and anarchistic users who, on a daily basis, use the modern information and communication technologies, such as collaborative networks for sharing ideas, opinions, data files, video, music, e.g. YouTube, MySpace, Hi5, FaceBook and others. Various media and technologies are being used for this purpose e.g. blogs, podcasts, P2P sharing, and highly personalised websites.

2. LENVIS NETWORK

2.1 Introduction

The lenvis project aims at research and development of an architecture and technical solution for integration of services, linking water, air and health domains, providing information through a user centric decision support network.
Through the development and integration of novel ICT solution it combines (Figure 1):

- SOA architecture, enabling collaborative services and a more efficient fusion of environmental data and modelling services and health risk assessment indicators. These components are termed as Service Provider Peers (SPP);
- Collaborative environment, facilitating collaboration between the end-users, in particular the EU citizens. These components are termed as User Peers (UP);
- The network management component, in order to assure the quality and security of the data and information provision on the lenvis network. These components are termed as Network Management Peers (NMP).

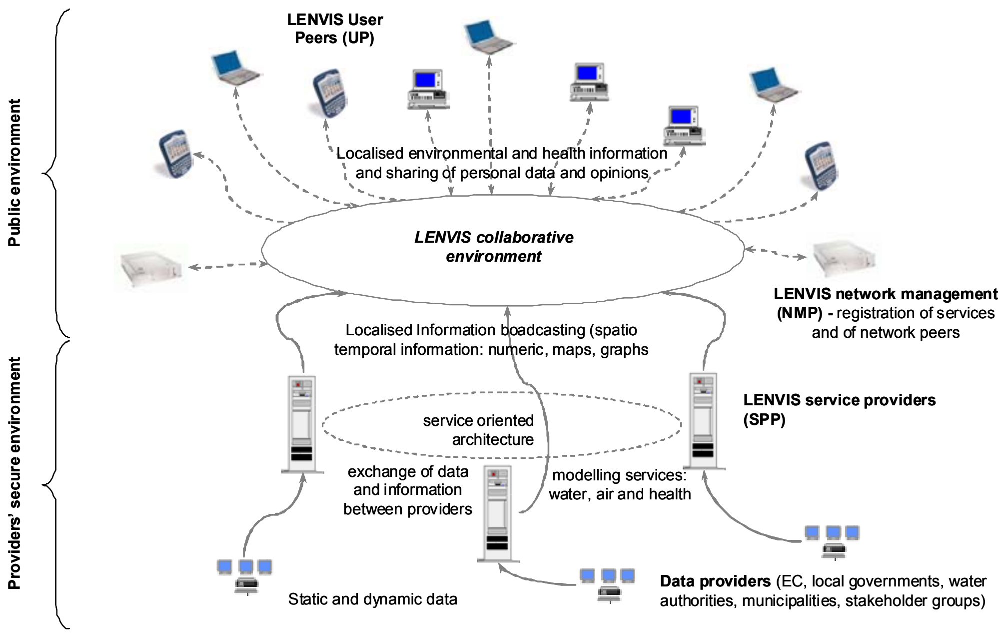


Figure 1. Schematic representation of the lenvis concept.

2.2 Service Provider Peers

The SPP is a generic environmental service provider. Each SPP provide one or more services and the interface to the service bus. A service repository contains the descriptions of the services. Environmental service providers register their services in this repository and service consumers access the repository to discover the services being provided. All SPPs cooperate according to a SOA approach. The SPP data providers can provide quality labelled data, connections to sensor networks and telemetry data, or they can combine services from other lenvis and non lenvis service providers (e.g. meteorological and geo services, Google map services, etc.).
Lenvis SPPs will also provide domain related modelling services, in the framework of the SOA architecture. Advanced environmental modelling systems for a regional scale as well as a local scale are used, e.g. MM5, CHIMERE, MOHID, and HydroNET. These systems are used together with health modelling algorithms to be developed and will be available as services in lenvis for an integrated assessment of air quality, water quality and the related health impact.

2.3 User Peers and Network Management Peers

The User Peers (UP) create the basis of the collaborative network. The main goal of this socio-technical network is to facilitate exchange of environmental and health-related data, information, feedback, analysis, models, documents, multimedia and other information between the stakeholders and EU citizens on a certain geographical location and about certain catchments, urban area or ecosystem. The UPs will run as a lenvis desktop application, a web portal and as light clients on semantic-enabled mobile devices (e.g. PDAs, cell phones, laptops). Key functionalities of the UPs are:

- graphical display of geo-temporal data in thematic maps: water quality for bathing, air quality levels health risk indicators e.g. for asthma patients.
- sharing of data and information, feedback and opinions;
- initiating polls and public opinions;
- collaboration facilities with other peers on the network.

The NMP provides the link between the first two components. The key real-time functionalities of NMPs are registration of all peers on the network, registration of available services on the network, user profiles management, management of meta information (dynamic) and routing of users.

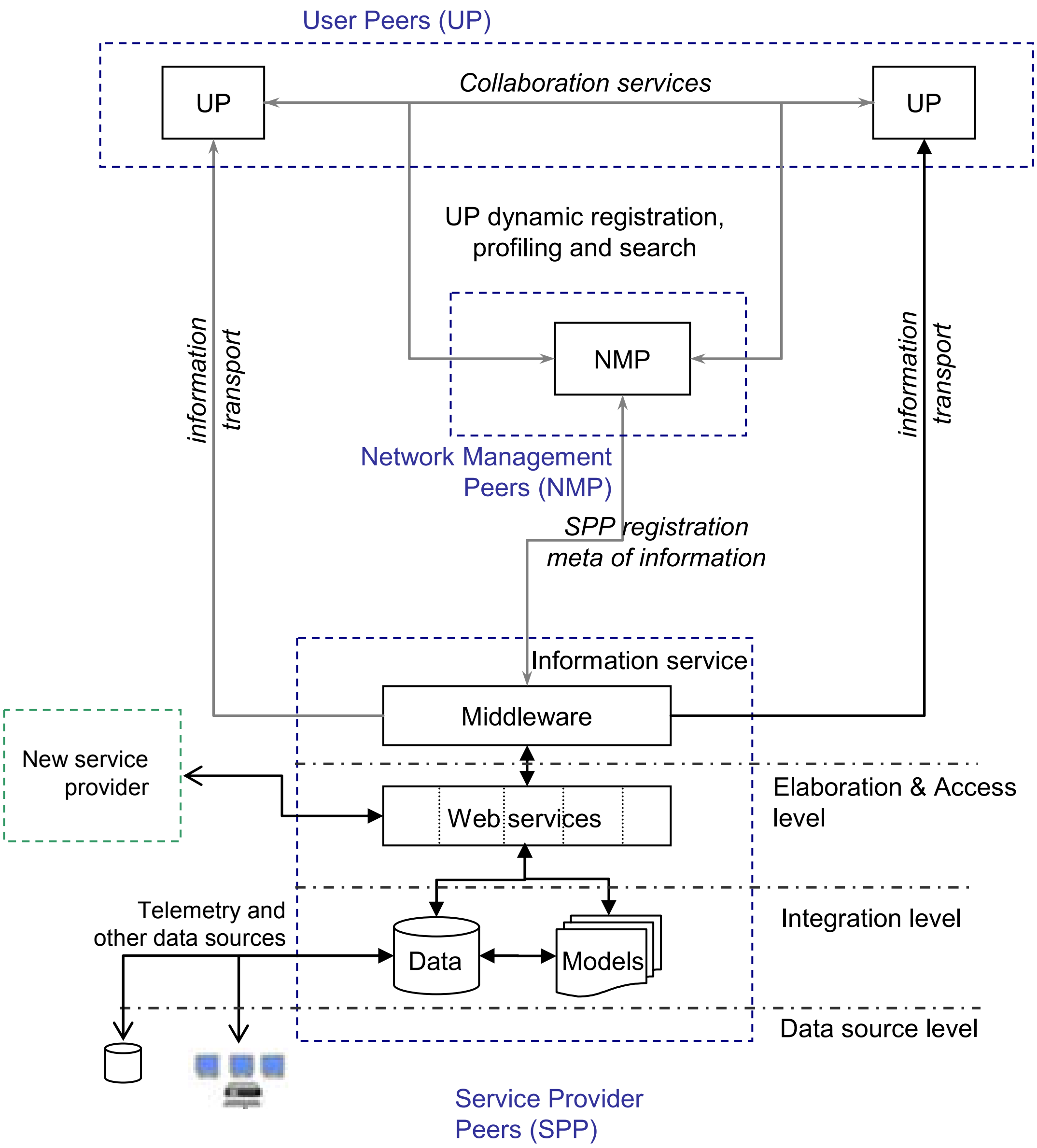


Figure 2. Logical architecture of the lenvis system.

3. CASE STUDIES

Lenvis, being an ICT project focused on developing novel applications for environmental management, will be validated in three case studies in those areas where epidemiological studies are already available: air pollution in urban areas (cardiovascular, respiratory diseases) and water pollution in rivers and estuarine areas (gastrointestinal and respiratory diseases).

3.1 Italy: Air and health

The objective for the case study in Italy is to use air quality monitoring and prediction and health services for the urban areas of Milan and Bari to analyse relationships between air pollution and health risks, to evaluate air quality management strategies for their effectiveness, and to forecast air quality conditions and probable effect on health. As air pollution is becoming a ‘hot topic’ among citizens and especially among young people (school and media are dedicating more and more space to this theme) a further goal of the case studies in Italy is to experiment how the availability of information will increase the awareness on this critical issue. The end-users are the University of Milan, the Municipality of Bari and the citizens of these urban areas.

3.2 Portugal: Water and health

The Portuguese case study area is the Lisbon coastal region, which presents extensive bathing areas in both the north and south banks separated by the Tagus estuary. All beaches are usually good for bathing and frequently visited by the inhabitants of Lisbon and foreign tourists, especially during the summer months. In the last decade there has been a great effort in the identification and eradication of point sources along the coast, but diffuse sources of faecal contamination associated to watersheds remain a problem. As a consequence, occasionally faecal pollution of the bathing waters posses a health risk. Since it is impossible to control all diffuse sources of faecal pollution in heavily populated areas (context for the Bathing Water Directive 2006/7/EC) an information system needs to be established that is able to alert the public and authorities for water quality problems. Lenvis will contribute to this objective addressing the current lack of integrated water quality predictive capacity and appropriate data management and dissemination systems by applying modelling services on existing watershed models and meteorological models. The lenvis system will be used to manage all the real-time monitoring and modelling data streams and to interact with the end-users on the basis of health risk alerts. The end-users are the water utility company (Sanest), the Portuguese National Health Service, and the visitors / recreational users of the beaches.

3.3 The Netherlands: Water, air and health.

The Dutch case study takes place in the Province of Noord-Brabant The province counts almost 2.5 million inhabitants, of which by far the most live in urban areas. In between the cities, there are numerous recreational lakes that are used intensively during summer. Both the air quality in the province and the bathing water quality of the lakes sometimes drops below health risk thresholds. The bathing water quality is mainly threatened by algae and external inflow of polluted water. The algae concentration is monitored by a number of point measurements in some of the lakes. The monitoring of water quality of the connected canals and streams is limited and the progression towards the bathing water is not modelled. High waters impose risks to the area as well as may be damaging the existing natures. Flood damage prevention projects have been carried out to research this. Alerts on high waters are the responsibility of the local waterboards, which take care of water quantity and water quality control. The air quality has strongly improved over de past 10 years, but still not all concentrations of substances are meeting the European thresholds. An air quality monitoring and information system is proposed to help to decide on measures to further improve the air quality, and a meteorological forecasting system to provide longer warning times in case of calamities. Lenvis addresses the need for modelling of both water quantity/quality and air quality and the subsequent health risks to provide timely warnings to the end-users in these three domains. HydroNET software provides the required data services for handling meteorological data and forecasts. Online modelling services will be used for hydrological, hydrodynamic and water quality modelling. The end-users are: the Province of Noord-Brabant, the waterboards, representatives of stakeholder groups (nature, recreation, etc) and the citizens.

4. VALIDATION AND TESTING

The validation and testing of the lenvis DSN is divided into two programmes. The first is the technical, functional testing of the system. The second is the end-user validation of the system. In the functional testing, the data and modelling service providers are involved.

Each component, e.g. real-time data delivery and environmental and health models, are first tested separately by the individual provider for each case study, and then in the integrated system. The functional testing of the lenvis DSN is lead by the individual developers of services and is also performed for each case study.

From this point onward the end-user validation starts. Here still functional properties will be incorporated, such as reliability and robustness of the system, however, emphasis is on getting the end-users experiences when using the lenvis DSN in (near) real-life situation testing. Different aspects such as accuracy of the information, availability of the information, relevance of the information, interpretability of the information and user friendliness of the system (software and hardware) will be assessed. Small dedicated user groups are being selected. They will include persons from organisations involved as end users, and a number of citizens. The citizens will interactively report, through a predefined format, on how the warnings through the system (would have) affected their actions.

5. ACKNOWLEDGEMENTS

The lenvis project is part of the FP7 ICT - Information and Communication Technologies Work Program, Challenge 6: ICT for Mobility, Environmental Sustainability and Energy Efficiency.
Involved partners are: HydroLogic (NL), UNESCO-IHE (NL), Project Automation (IT), Technical University of Lisbon - IST (PT), Aria Technologies (FR), EsaProjekt (PL), University of Milano Bicocca (IT), Hidromod (PT), Province of Noord-Brabant (NL) and Municipality of Bari (IT). More information on this collaborative project as well as the partners involved can be found on: www.lenvis.eu.

6. CONCLUSIONS

The main goal of the lenvis project is to develop an innovative collaborative decision support network for exchange of location-based environmental and health services between stakeholders. This is done by developing a generic ICT solution that combines service-oriented-architecture (SOA) and user-centric approach (collaborative network). Air and water data and models are integrated in a SOA architecture, to be available as services for environmental management. These will provide meteorological, air and water quality information at specific geographical location. Health indicators are included on top of these data streams. Lenvis aims at supporting governments in communicating environmental and health information with the general public. In particular it addresses the young generation Y, by entering popular digital social networks and providing the information in popular interfaces such as collaborative websites and on mobile devices such as mobile phones and PDAs.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Empowering Environmental Knowledge Convergence, Harmonisation and Integration: Challenges and Approaches

Pier Giorgio Marchetti
European Space Agency (pier.giorgio.marchetti@esa.int)

Abstract: This paper builds on experience made at ESA addressing these issues in global programmes and within the GIGAS Support Action. It outlines specific proposals addressing the methodologies in order to empower the semantic and technical interoperability across environmental information systems and services.

Keywords: Earth observation; standardisation; architecture and interface design; convergence and harmonisation strategy; knowledge sharing; information discovery

1. INTRODUCTION

Efforts in addressing global issues related to the monitoring and management of environmental global change phenomena are supported by environmental observations such as the ones to be provided by the GMES programme. The earth observation and in-situ data provided to the GMES services, together with the spatial information made available by the INSPIRE directive should provide seamless information and services in order to ensure that the objective of environmental sustainability can be achieved.

The availability of leading edge technologies empowers new information flows, at the same time challenges the user orientation in the wealth of available information. The concurrent development of initiatives requires the convergence, harmonisation and standardization in order to ensure a sustainable development.

GMES is also the European contribution to the worldwide monitoring and management of the environment within the Group on Earth Observations (GEO). GEO was established, with the goal of addressing the information requirement for the environment on a global scale to be achieved via a 10 year implementation plan of an integrated Global Earth Observation System of Systems (GEOSS). These parallel initiatives demand a high level of consistency and convergence. The GIGAS Support Action financed by the Directorate General Information Society of the European Commission. GIGAS (GEOSS, INSPIRE and GMES, an Action in Support) addresses methodologies for the convergence of the GMES, INSPIRE and GEOSS initiatives (see http://www.thegigasforum.eu).

The access to earth observation data for the extraction of the relevant information may represent a significant part of the effort in the data exploitation and service delivery. To address these issues a harmonization effort was launched in order to align the interface between the space segments, known as HMA project, which has provided the standards for the interfaces within the GMES Space Component.

Large scale interoperability projects like HMA, and initiatives like INSPIRE and GEOSS define architecture and interoperable interfaces making use of the Reference Model of Open Distributed Processing - RM-ODP [1998], Uslander, [2007]. This paper addresses the approach taken in re-using the same methodology in the context of the analysis, comparison, convergence and harmonization of different architectures, specifically addressing the convergence elements with particular concern for the semantic interoperability.

2. HMA INTEROPERABILITY SCENARIO/USE CASE DEFINITION METHODOLOGY

The following interoperability scenario and use case definition methodology has been successfully and effectively used in HMA and can be reused in other interoperability – harmonisation - standardisation related scenarios.

2.1. Use Cases

- The interoperability - harmonisation context is modeled through a set of use cases defining the system, the subcomponents and the external interfaces.
- Use Cases are broken down into scenarios describing the interactions across "centres" (or Ground Segments) performing "services" and external interfaces.

2.2. Scenarios

- Scenarios describe Interactions with UML-like sequence diagrams.
- A Pareto-like approach is used addressing common scenarios (20%) which can be shared by most of the stakeholders (80%).
- Common Scenarios shall not retain the identification of the "owner" or the initiator.
- System interfaces shall be abstracted to 2-3 high levels across centres.
- System and implementation specific aspects are out of scope.
- Only scenarios which reach consensus are formalized.

2.3. Requirements

- High level requirements shall be derived from each scenario
- High level requirements shall be associated with the high level services and the centre or ground segment - GS
- The distribution of roles across centres and/or GS shall be limited to what necessary for the definition of interfaces
- High level requirements shall be traced into Service Interface Requirements

2.4. Architecture

- High level requirements and Service Interface Requirements shall be traced to the architecture
- Architecture shall trace the service interface control documents - ICDs
- A RM-ODP approach is used

The above process may be applied recursively at different level of definition and focus.
Scenarios at system level (i.e. where the system is shown as a single black-box entity), may be expanded into lower lever scenarios. These lower level scenarios may contain a breakdown of the system into the main services or subcomponents, showing internal system interactions and identifying more precisely the external interfaces. These lower level scenarios put the focus on the system sub-components.
With the same mechanism, if the system level scenarios help to identify system level requirements (i.e. “the system shall..”) then the lower level scenarios lead to sub component level scenarios (i.e. “Sub-component X shall..”, “Sub component Y shall..”).

3. THE GIGAS PROPOSED HARMONIZATION APPROACH

The approach proposed by GIGAS (see Biancalana et al. [2008]) is based on three elements:

- Technology watch
- Comparative analysis
- Shaping of initiatives and standards

The above steps will be repeated twice (in two loop) during GIGAS lifetime but can be repeated multiple times after the project conclusion if a persisting infrastructure is established. This paper concentrates on the methodology for technology watch and comparative analysis.

3.1. Technology Watch

The complexity of the GIGAS scenario involving huge systems (i.e. GEOSS, INSPIRE, GMES etc.) entails the interaction with different heterogeneous partners, each with a specific competence, expertise and know-how.

The methodology proposed is based on an RM-ODP based study supported by interoperability use cases and scenarios used to derive requirements.

GIGAS will monitor

- the INSPIRE, GMES and GEOSS evolution and analyze the requirements, the standards, the services and the architecture, the models, the processes and the consensus mechanisms with the same elements of the other systems under analysis;
- activities in the fields of standard development that are part of the three initiatives; this task will provide the basis for how these three initiatives will strategically support consensus and efficient standards development going forward;
- architecture, specifications, innovative concepts and software developments of past or ongoing FP6/FP7 research topics.

The use of an RM-ODP approach is selected as:

- most of the architectural approaches to be compared are based on RM-ODP,
- it supports distributed processing,
- it aims at fostering interoperability across heterogeneous systems,
- it tries to hide distribution to systems developers.

However, as most of the systems to be considered have the characteristic of a loosely-coupled network of systems and services instead of a “distributed processing system based on interacting objects”, the RM-ODP concepts are tailored for the GIGAS needs.

The usage of RM-ODP for GIGAS Requirements and Technology Watch is two-fold:

- Architectural analysis: It is performed for all projects and initiatives. Its purpose is to identify possibilities but also major obstacles for interoperability. Furthermore, it identifies the major use cases to be analysed in more detail.
- Use Case Implementation Analysis: It is used to describe how selected use cases of the projects and initiatives are implemented in the different architectures. Its purpose is to identify technological gaps and concrete problems of interoperability. It is performed only for selected use cases.

The output of the Technology Watch is an RM-ODP based report containing parallel analysis on the same aspects on the three initiatives integrated by analysis of relevant FP6-FP7 projects and standardization activities.

3.2. Comparative Analysis

Based on the outcomes of the previous monitoring tasks, GIGAS undertakes a comparative analysis on solutions, requirements, architecture, models, processes and consensus mechanisms used by INSPIRE, GMES and GEOSS, taking into account the inputs from the monitoring of FP6/FP7 research projects and the ongoing standardization activities.
Initiative Contact Points will insure that the overall policy framework and schedules for each of the three initiatives will be factored in.
The result of the Comparative Analysis includes:

- A list of recommendations to GEOSS, INSPIRE and GMES to be expanded and processed in depth in the following shaping phase
- The identification of technological gaps to be explored in the following shaping phase.
- Guidelines and objectives for the architectural approach within GIGAS
- Analysis on the schedules of the three initiatives and on the FP6/FP7 programs and standardization activities, with identification of key milestones or intervention points.

4. GIGAS CONSENSUS BUILDING PROCESS

Consensus Building is defined as "an array of practices in which stakeholders, selected to represent different interests, come together for face-to-face, long term dialogue to address a policy issue of common concern" Innes and Booher, [1999]. In this approach, consensus building can lead to the needed convergence and agreements, as well as to long term changes among the stakeholders.
Therefore a good consensus building process:

- Includes representatives of all relevant stakeholders
- Involves experts helping to identify possible solutions;
- Is driven by an agreed purpose;
- Is open, allowing participants to decide on working rules, objectives, tasks, priorities, and discussion topics;
- Engages participants and facilitates learning through in-depth discussion;
- Encourages emerging of new challenges and creative thinking;
- Incorporates high quality information and assures agreement on its meaning;
- Achieves consensus only after discussions have fully explored the issues and interests.

The methodology proposed for GIGAS by Max Craglia [2008] builds on six steps Susskind et al., [1999]:

1. convening, which means identifying: stakeholders and their representatives, priorities and potential areas of disagreement, convening the meeting(s)
2. clarifying responsibilities, defining an agreed convergence process, its timeline and the participants' roles
3. deliberating, which entails brainstorming, "lateral thinking" and finally identifying win-win solutions
4. deciding, that is aiming at an agreement that maximizes benefits for all
5. implementing agreements, which relies on the effectiveness of the process and the sense of ownership of the agreement reached by the stakeholders
6. organizational learning, that is establishing a consensus building process that enables the growth and evolution of the community

5. KNOWLEDGE EMPOWERMENT

The concept of semantic interoperability addressed above demands work to be performed in several different directions: the interoperability across conceptual data models and information models as well as knowledge empowerment and sharing. Taxonomies and classifications have been instruments for the sharing of the knowledge. In computer science ontologies have been used as knowledge representation. The term ontology is used to mean a specification of a conceptualization according to Gruber [1993]. The dynamic link of ontology and discovery empowers knowledge sharing and facilitates discovery. This is an essential element to consolidate domain knowledge but as well to establish the bridge between the layman's and specialists view's, and the conceptualization, naming and perception as depicted in the "Montezuma and the horses" tale in Eco [1999].

If we consider the case of Earth observation products the work on ontology and terminology shall allow to use terms which are not domain specific to identify products, services and processing components. The specific ontology shall be as simple as possible, address multiple domains, show limited dependencies from evolution and changes and allow multilingualism. A proper open discovery mechanism should allow a dynamic linking of the GEMET thesaurus and other multi-domain thesauri, see Felluga [1995].

6. CONCLUSIONS

This paper presents an overview of methodologies applied to large scale earth observation projects. The following elements summarize best practices supporting the empowerment of environmental knowledge:

- Establish the technical terms for the harmonization
- Build consensus processes
- Empower knowledge sharing with ontology and discovery

ACKNOWLEDGEMENTS

The authors wish to thank the colleagues within the GIGAS consortium for their input and in particular Andrea Biancalana (Elsag-Datamat) and Sergio D'Elia (ESA) for their comments in the review of the paper.

REFERENCES

Biancalana, A., Coene, Y., Marchetti, P.G, Uslander, T., GIGAS Methodology for Technology Watch, *GIGAS Project Deliverable*, 2008

Craglia, M., GIGAS Consensus Building Methodology, *GIGAS Project Deliverable*, 2008

Eco, U., Kant and the Platypus, Secker & Warburg, London, 1999

Felluga, B. (Ed.), Multilingual Thesaurus for the Environment. Classification Scheme, CNR, Roma, 1995

Gruber, T., Toward Principles for the Design of Ontologies Used for Knowledge Sharing, *International Journal Human-Computer Studies 43*, 907-928, 1993

Innes J., Booher D., Consensus Building and Complex Adaptive Systems, *Journal of American Planning Association*, Vol. 65 (4), 412-423, 1999.

Marchetti, P.G., Biancalana A., eds., Feasibility Study CEN 202 Space Standardisation, EO, GMES and Dual Use - State of Play, EOEP-INPR-EOPG-TN-2008-0001, 2008

Marchetti, P.G., Exploiting Earth Observation Missions: Opportunities and Issues in Ground Segment Interfaces Harmonization, *NATO Advanced Research Workshop Geographical Information Processing and Visual Analytics for Environmental Security*, Springer, 2009

RM-ODP Reference Model of Open Distributed Processing - ISO/IEC 10746-1:1998, 1998

Susskind L., McKearnon S. and Carpenter S., Consensus Building Handbook, Sage Publications, 1999

Uslander, T., edr., Reference Model for the ORCHESTRA Architecture, OGC 07-97, 2007

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

An Holistic View of Coverage Model and Services for SISE-SEIS

S. Nativi[a] **and A. Woolf**[b]

[a] *Italian National Research Council - IMAA, and Univ. Florence at Prato, Prato IT 59100, Italy (stefano.nativi@cnr.it)*

[b] *STFC Rutherford Appleton Laboratory, Oxon OX11 0QX, UK (andrew.woolf@stfc.ac.uk)*

Abstract: There has recently been a growing recognition of the need for interoperability – not only in traditionally 'geographic' disciplines (e.g. mapping, physical geography) but also in more science-focussed disciplines like Earth systems science, geology, meteorology, oceanography. These latter application domains bring a fresh viewpoint to the nature of spatial information. While the conventional geographic disciplines have a primary focus on a 'feature' view (points/lines/polygons with thematic attributes), the scientific domains utilise a conventional 'coverage' view (thematic properties distributed over some region in space and/or time). For integrated information systems like SEIS and SISE, however, a more harmonised model is required. Fundamentally, the feature- and coverage-views are complementary, and any general framework must support both viewpoints. A framework for harmonisation is offered through the observation process that lies at the heart of much Earth science data. By recognising that an observation process samples the natural environment (which may be represented through a feature view) and generates a coverage result, we are able to reconcile these hitherto different approaches. This harmonized model supports the different types of discovery and access services useful to serve the heterogeneous SEIS use cases.

Keywords: SDI; Coverage model; Observation & Measurement model, Feature model.

1. INTRODUCTION

The growing area of environmental informatics is concerned with providing integrated access to a range of advanced information and processing resources for the environment. Both the US and European premier scientific unions are recognising this – the American Geophysical Union (AGU)[1] has established an Earth and Space Sciences Informatics (ESSI) Focus group, and the European Geosciences Union (EGU)[2] has recently created a new scientific Division for ESSI. Primarily, interoperability and metadata are identified as two key technologies in discovering and enabling access to usable information and processing resources for the environment. Interoperability is achieved by adopting a Service-Oriented Architecture (SOA) approach and applying international standards, as far as service interfaces and data models are concerned. International standard organizations,

[1] http://www.agu.org/

[2] http://www.egu.eu/

such as ISO TC211[3] and OGC[4], have been working to develop such a standards baseline for the geospatial information domain. This approach has been supported and adopted by important Spatial Data Infrastructure (SDI) initiatives such as the European INSPIRE[5].

Earth science is at the forefront of applying advanced computing technology to the solution of pressing environmental problems. GMES and GEOSS initiatives represent a couple of significant cases in point. In fact, important environmental problems demand an integrated modelling of coupled physical processes, global datasets, and a multi-disciplinary coordinated approach (e.g. biologists working together with climatologists to determine the impact of warming on species distributions).

However, the information modelling approach used by the SDI initiatives and GMES and GEOSS is novel – for the most part – to the environmental and earth science communities, and there are significant challenges in bridging the conceptual gap. In earth science there have been some applications of this modelling approach – notably in the geosciences and atmosphere/ocean. A valuable example is represented by the OGC GALEON IE (Geo-interface for Air, Land, Earth, Oceans NetCDF Interoperability Experiment)[6] which deals with specifying and using standard interfaces to foster interoperability between data systems used by the traditional GIS community and those in the community referred to as the Fluid Earth Sciences (FES, mainly oceanography and atmospheric science). These attempts to bridge legacy information environments in earth science with SDI have both proven the general feasibility and identified weaknesses and challenges.

Holistic interdisciplinary approaches and lack of common data models and semantics are important research challenges to be addressed in order to achieve a Single Information Space for the Environment (SISE) in Europe [Juceviciene, 2008].

Geospatial information models for interoperability recognize three important concepts: feature, coverage and the more general map. ISO TC211 introduced two fundamental concepts to map both discrete and continuous real world phenomena: features and coverages. For Earth Sciences, a ‘coverage’ or field view of information is very predominant – much earth science data is regarded as a field over some region of space and/or time, rather than a discrete spatial object with attributes. Multiple coverage types exist; mainly, they are disciplinary related. Virtually, any geospatial data may be viewed as an instance of a coverage type. Different coverage types are characterized by different coverage domains and coverage functions.

There exists a clear need to develop an holistic approach to model, discover, access and use environmental coverage data types. For SEIS-SISE, it is particularly important to rely on effective and flexible models and service interfaces for coverage datasets. This holistic approach must be clearly harmonized with the General Feature Model (GFM) adopted by the international standardization frameworks for geo-information –e.g. ISO and OGC.

The manuscript is structured as follows. The first section briefly discusses the coverage concept as introduced and used by standard data models for geo-information, outlining the different coverage types that characterize important geospatial communities. Having in mind these different views, the second section discusses an holistic coverage and feature model for SEIS-SISE and the need for one or more types of coverage access service. The final section summarizes the manuscript conclusions.

2. THE COVERAGE CONCEPT

The coverage concept was defined to summarise the different conceptual and physical representations of a traditional image, going further by enlarging the spectrum of geospatial information that can be represented this way. Hence, the “coverage” term refers to any data

[3] http://www.isotc211.org/

[4] http://www.opengeospatial.org/

[5] http://inspire.jrc.ec.europa.eu/

[6] https://sites.google.com/site/galeonteam/Home

representation that assigns values directly to spatial position. Thus, a coverage may be seen as a function from domain –commonly a spatial, temporal or spatiotemporal domain– to an attribute range. A coverage associates a position within its domain to a record of values of defined data types. According to ISO TC211, coverage is a subtype of feature that has multiple values for each attribute type, where each direct position within the geometric representation of the feature has a single value for each attribute type [ISO 19123].

Therefore, a coverage may be generally modeled through: a domain (the well-defined set of direct position the coverage deals with); a co-domain or attributes range; a coverage function (the rule that associates each element from the domain to a unique element in the co-domain).

2.1 Coverage Types

By way of example, we consider a number of coverage types that are of widespread interest within the oceanographic and meteorological communities.

As with many earth science disciplines, these communities collect data through observational campaigns or deployment of monitoring instruments. Both the measured values and the location/time at which they are measured are important – fundamentally this represents a coverage view. Moreover, it is usual for practitioners to classify coverage types into classes based on the geometry and topology of the discrete coverage domain.

There are very good scientific reasons why this should be so. Physical processes occur in the natural world across a wide spectrum of spatial and temporal scales, and considerable science informs the design of experimental sampling strategies. Conversely, the geometry and topology of observation sets are a fundamental determinant of the scientific uses to which they may be put. Moreover, the properties of the instruments used to generate data themselves place constraints on their interpretation (e.g. as regards accuracy, precision, calibration, required post-processing, etc.). These two factors – the scientific utility of a sampling regime, and the limitations of an observing process – lead to a natural, scientifically important, classification of data types along these axes. Quite often the two are highly correlated (certain instruments generate certain samplings), and so scientific communities of practice adopt more abstract conceptual information classes that nevertheless reflect artifacts of sampling or instrument-type. Within the meteorological and oceanographic communities, broad information classes based on measurement-set geometry and topology have almost universal acceptance. We next describe some examples:

For example, the US National Oceanographic and Atmospheric Administration (NOAA) is developing a plan for a Global Earth Observing Integrated Data Environment (GEO-IDE) to integrate measurements, data and products and create interoperability across data management systems. In the GEO-IDE Concept of Operations[7], the following 'structural data types' are defined: *Grids, Moving-sensor multidimensional fields, Time series, Profiles, Trajectories, Geospatial Framework Data, Point Data, Metadata*.

The ESRI 'ArcMarine' Data Model for marine data includes classes like *Instant, Location Series, Time Series, Profile Line, Track, Sounding, Survey, {Ir}Regularly Interpolated Surfaces, Mesh Volume, etc*. File formats such as netCDF and NASA Ames utilize data models that reflect these structures (e.g. netCDF four-dimensional gridded lat-lon-height-time variables, or NASA Ames time-series at a point). The netCDF Common Data Model (CDM) and the Climate Science Modelling Language (CSML) adopt very similar classifications, Table 1.

3. AN HOLISTIC APPROACH

In an object-oriented approach every business entity is an object type. Analogously, in the ISO feature-based approach every business entity is a feature type. Actually, in this case

[7] https://www.nosc.noaa.gov/dmc/docs/NOAA_GEO-IDE_CONOPS-v3-3.pdf

they are geographical features: an abstraction of a real-world phenomenon that is associated with a location relative to the Earth [ISO 19107].

Therefore, in the ISO GFM framework a coverage is simply harmonized treating it as a feature sub-type; this is often implemented by defining “logic” features rather than physical ones. The sub-typing connection is quite general. Hence, there is a clear need to investigate such association in order to outline an holistic approach for implementing a more effective and flexible harmonization model. The next chapters will investigate this issue proposing a model for an harmonization solution.

Table 1: Comparison of CDM and CSML v2 feature types [Caron, 2008]

CSML Feature Type	CDM Feature Type
PointFeature	PointFeature
PointSeriesFeature	StationFeature
TrajectoryFeature	TrajectoryFeature
PointCollectionFeature	StationFeature at fixed time
ProfileFeature	ProfileFeature
ProfileSeriesFeature	StationProfileFeature at one location and fixed vertical levels
RaggedProfileSeriesFeature	StationProfileFeature at one location
SectionFeature	SectionFeature with fixed number of vertical levels
RaggedSectionFeature	SectionFeature
ScanningRadarFeature	RadialFeature
GridFeature	GridFeature at a single time
GridSeriesFeature	GridFeature
SwathFeature	SwathFeature

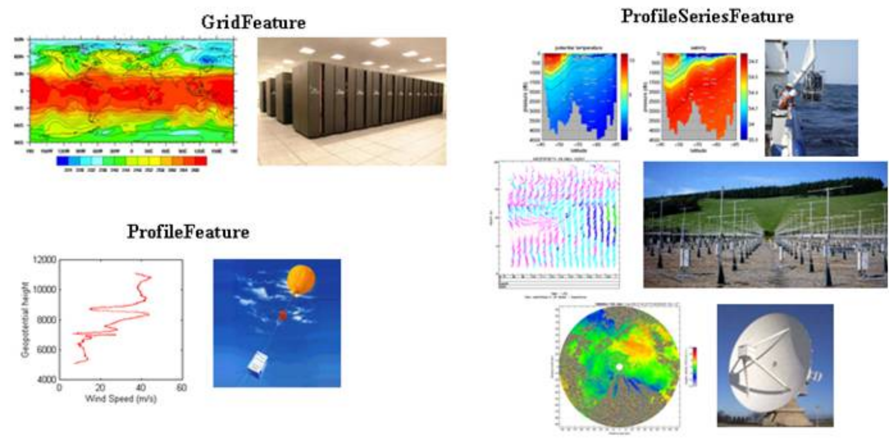


Figure 1: Illustration of some coverage types

3.1 The General Conceptual View for Geo-information

Geoinformation is used to describe objects, phenomena or processes related to the Earth [Molenaar, 1984]. This is commonly done in the form of state description at a certain moment or, for processes, as a series of state descriptions. These may refer to one or more aspects characterizing terrain objects (geographical features). These aspects are given as thematic attributes in relation to object shapes (geometric data). In the most straightforward form the thematic attributes are directly linked with positional data –in a given N-dimensional domain. Hence, the positional data serve as a vehicle to link different types of thematic data or to link data obtained at different moments. A higher information level can be obtained by the introduction of terrain objects (i.e. features class instances) [Molenaar, 1984]. In fact, thematic attributes are not linked directly to the positional data, but to the terrain objects (feature instances).

These two approaches reproduce two traditional and different space conceptualizations: the object and the field views (Couclelis, 1992; Goodchild, 1992; Peuquet, 1984). The object approach –sometimes called as geo-relational, boundary or “bird-eye view”– considers space as being ‘empty’ and populated with discrete entities embedded in space, while the field approach –sometimes called as tessellation, direct-position or “worm-eye view” – considers the space as being continuous, and every location in space has a certain property (Ledoux, 2008).

The ISO GFM approach adopts the object/geo-relation approach, focusing on feature classes and instances definition and description; geometric data are structured using feature types (i.e. geometric objects). While, the ISO coverage realm refers to the field/direct-position approach, focusing on the definition and description of coverage domains and mapping functions (i.e. coverage functions). Tessellation or polygonal mesh types are introduced to manage data structures –realizing commonly used implicit geometries, such as regular grids. Figure 2 depicts a general harmonization model for the two approaches.

Geographical features are characterized by a geometry (i.e. shape) and a thematic description. The feature sets defined through the geometric characteristics are called “feature types”; while, the sets defined through the thematic characteristics are called “feature classes”.

Locations of a field domain are mapped to thematic attributes applying a coverage-function. A domain is comprised of a set of direct positions. This set may be infinite – realizing continuous coverage; commonly, the set consists of a finite collection of points or geometric objects (e.g. tiles); they locate the samples or “ground truth” specimens of a field.

It is possible to find the spatial structure of a coverage thorough its positions topology. The topology is built up through the connectivity of neighboring domain elements (e.g. tiles). Neighboring coverage domain elements are connected within a feature instance if they have the same attribute values (Molenaar, 1991). Feature classes geometries are represented by connected domain elements –see Figure 2. This leads to “recognize” and extract feature class instances from coverages. The depicted schema supports also the other way around –i.e. to get coverage from feature class instances.

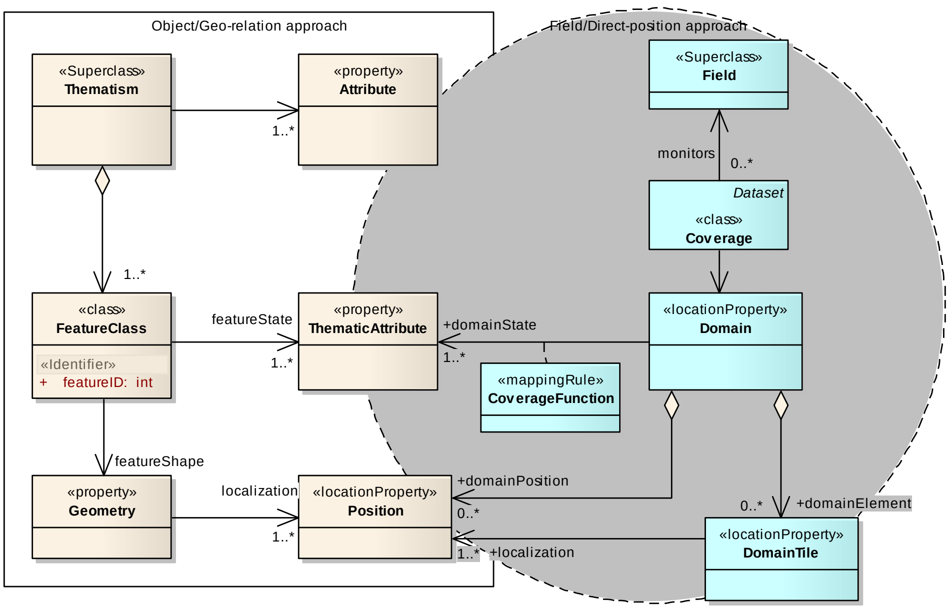


Figure 2. General approaches for Geo-information modelling

3.2 The harmonization context: a new view

In the SISE-SEIS context, an effective harmonization model can be motivated by considering the process of observing or measuring the environment. This is captured by the OGC/ISO Observations and Measurements (O&M) conceptual model [ISO 19156, 2008]. In natural language, the model states (see the dotted box in Figure 3) that an *Observation* is an action whose result is an estimate of the value of some *Property* of a *FeatureOfInterest* (FOI) obtained using a specified *Procedure.* Considering the general geo-information concepts previously discussed, the O&M model can be easily used to develop an harmonization framework for Environmental information systems. Figure 3 shows the context view of this harmonization model.

According to the O&M specification, *the FOI of an observation may be any feature having properties whose values are discovered by observation. In general, this will be of a type from catalogue representing the application domain for an investigation.* Normally, the FOI will be a so-called 'domain feature', representing an identifiable real-world spatial object. However, important for harmonizing feature and coverage views is to consider the case where the FOI exists only for the purpose of 'sampling' the physical environment (the *SamplingFeature* class of Figure 3). *SamplingFeature* is a FOI which may realize several observations concerning any identifiable feature. Examples include a weather balloon measuring temperature as it ascends through the atmosphere, or a moored tide-gauge measuring sea-level time-series at a location within a harbor. In those cases the observations sample a field (see Figure 1); thus, they are modeled through a *FieldSampling* class in Figure 3. In many of these cases, the result of the observation is a discrete Coverage (see Figure 1 and Figure 3) – a set of attribute values (e.g. temperature, sea-level) over a spatial, temporal, or spatiotemporal domain (e.g. the trajectory of a weather balloon, or the time instants of recorded sea-level). This realization provides the key step towards a model for harmonizing feature and coverage views.

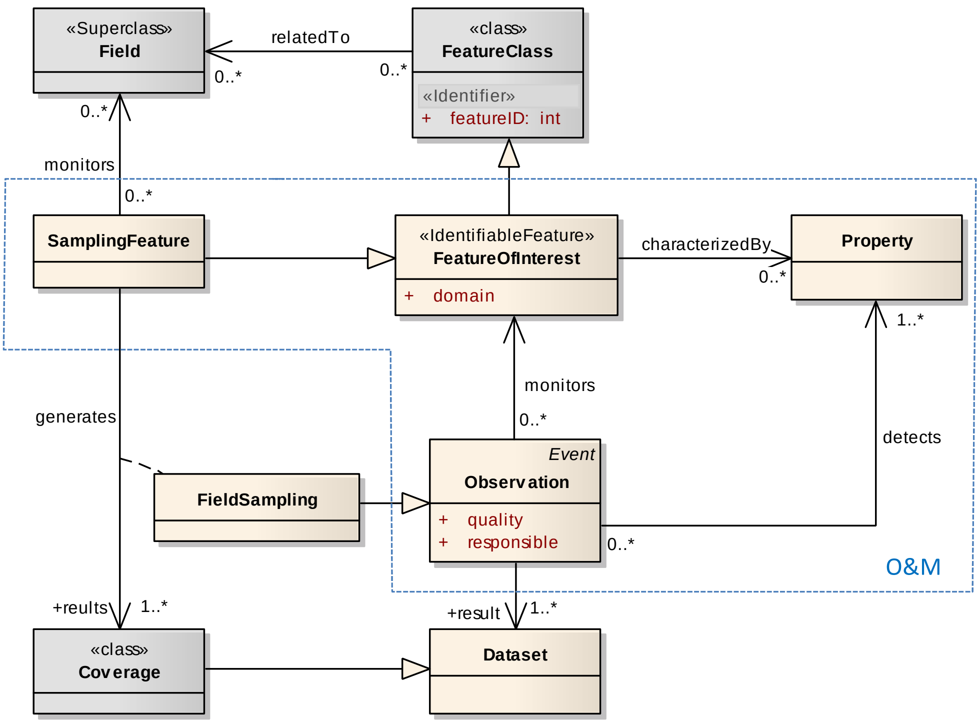


Figure 3 Harmonization model: context view

3.3 The Harmonization Model

We outlined above the key general pattern that recurs in many applications of O&M in environmental applications: the *FeatureOfInterest* is a *SamplingFeature*, and the result of the Observation is a Coverage. This realization provides a mechanism to integrate feature and coverage views through a harmonized model utilizing the O&M framework. Recalling the O&M model, an essential constraint is that the Observation result must be consistent with the observed property. For the case of a *SamplingFeature*, therefore, we propose that the observed property (*Phenomenon* class, Figure 4) is equivalent to the range of the Coverage result (or more strictly the semantic rangeType of the ISO 19123 CV_Coverage).In addition, we recognize that the *SamplingFeature* is intended to incorporate geometric aspects of the sampling regime, and propose that it should be equivalent to the domain geometry of the Coverage result (see again Figure 4). In the specific case of a coverage domain sampled through a set of tile elements, the survey procedure includes tessellation implementation information (*TesselationProcedure* call, Figure 4), such as: tile shape, geometry type, distribution, etc.

Thus, we arrive at a harmonization model for integrating feature and coverage views: there exists a real-world ‘ultimate’ FOI having properties that may be represented as continuous coverages. An example would be ‘The Atmosphere’ having a property ‘temperature’ that is a continuous coverage over a four-dimensional (x-y-z-t) domain. However, in practice we are limited to observing this ultimate domain feature only at discrete locations – the FOI in this case is not ‘The Atmosphere’ but rather a *SamplingFeature* that exists only to provide a concrete focus for the observation. The result of such an observation is a discrete Coverage (which may be classified according to broad classes of geometry and topology, as discussed in section 2.1 earlier). The domain geometry of this coverage is reflected in the geometry of the *SamplingFeature*, and the range of the coverage is an implied thematic property of the *SamplingFeature* (providing a consistent semantic closure for the observed property).

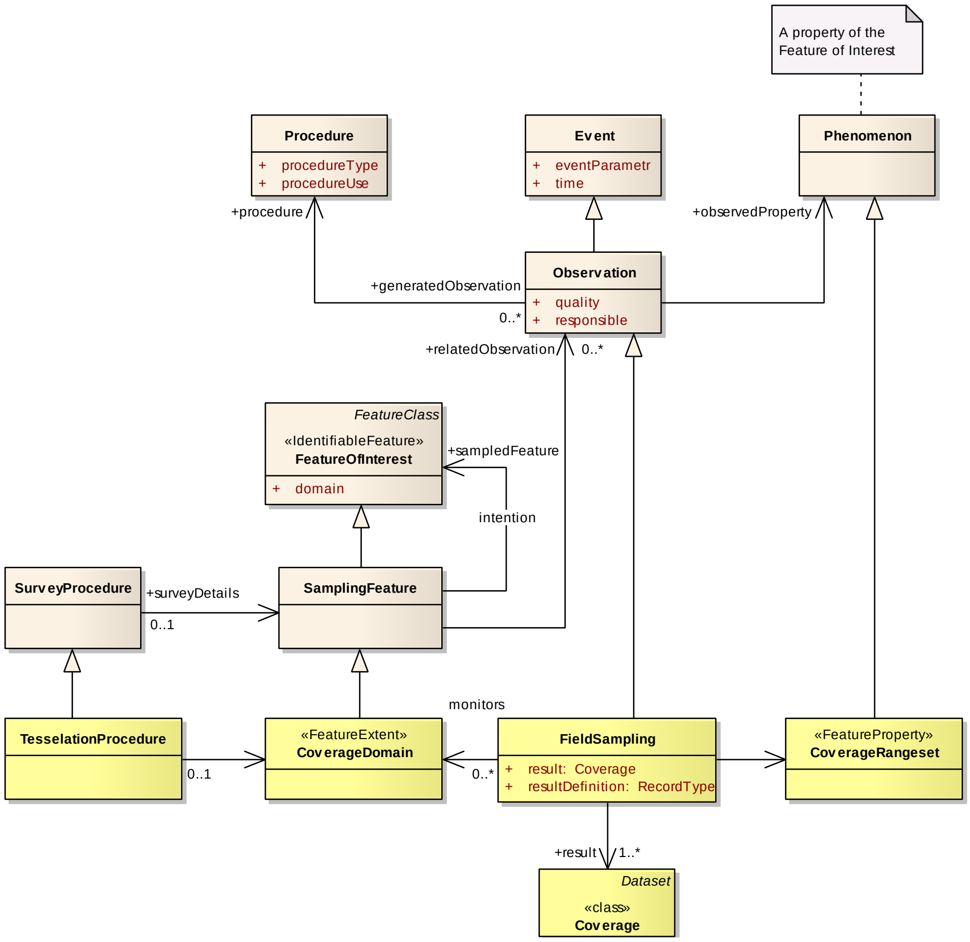


Figure 4 Harmonization model

3.2 Access services for SISE-SEIS

The harmonized model supports both the traditional feature based (e.g. FOI) and coverage based (e.g. sampled datasets) access services for an environmental information framework. For instance, standard service interfaces like the Sensor Observation Service or Web Feature Service may be more suited to the 'feature' view, while the Web Coverage Service may be more suited to the 'coverage' view. Besides, it enables the implementation of an advanced type of access services which make use of the traditional ones. These services provide the access to high level artifacts (e.g. phenomena, event, observation, procedures, etc.) avoiding to discriminate about the structure of their instances and representations. These two levels of services are extremely important to address the use cases heterogeneity characterizing complex and large "system of systems" such as SISE-SEIS.

4. CONCLUSIONS

For SISE-SEIS there exists a clear need to develop an holistic approach to model, discover, access and use environmental coverage data types. This holistic approach must be clearly harmonized with the ISO GFM. Coverage and GFM reproduce the traditional and different space conceptualizations which characterize geoinformation: the object and the field views.

We proposed a possible context to harmonize these two different views. This context is then implemented using the OGC/ISO O&M model. In fact, we recognize that in many cases with *SamplingFeatures* we get a Coverage result. Actually, the Coverage domain is the SamplingFeatures and the Coverage range is a feature property.

For a Shared Environmental Information System, these complementary views are important, and it is important also to realize that they do not conflict but rather represent alternative representations of the real world. A ‘feature view’ regards a feature class having thematic attributes that are themselves coverages, while a ‘coverage view’ considers just those attributes themselves. These two views are integrated through the act of observation: extensive properties or attributes of a real-world feature are sampled (by a ‘sampling feature’) leading to a discrete coverage result with the domain geometry equivalent to the sampling regime. A large number of environmental thematic areas (e.g. within Annex II and III of the INSPIRE European directive) utilize this dual feature-coverage view and would benefit from application of this harmonized model within their application schemas.

ACKNOWLEDGEMENTS

The authors wish to thank Paolo Mazzetti and Lorenzo Bigagli (from CNR-IMAA, Italy) for their kind collaboration.

REFERENCES

Barton, C., Fox, P., Khalsa, S.J.S., Jackson, I., Messerotti, M., Wyborn, L., Nativi, S., (ed.s), *Report of the ESI Summit 13-14 March 2008*, Rome Italy, June 2008.

Caron, J., “Common Data Model Scientific Feature Types”, *UNIDATA/UCAR presentation*, available at: http://conan.unidata.ucar.edu/Presentations/UPCsemseries/caronfeaturetalk.pdf, 2008.

Couclelis, H., People manipulate objects (but cultivate fields): Beyond the raster-vector debate in GIS. In AU Frank, I Campari, and U Formentini, editors, *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, volume 639 of Lecture Notes in Computer Science, pages 65–77. Springer-Verlag, 1992.

Goodchild, M.F., 1992, Geographical data modeling, *Computers & Geosciences*, 18(4):401–408.

ISO/CD 19156, Geographic information — Observations and measurements, *ISO/TC 211 document* N 2616, 2008.

ISO/FDIS 19123, Geographic information—Schema for coverage, geometry and functions, ISO/FDIS 19123:2005(E).

ISO/IS 19107, Geographic information—Spatial schema, ISO/IS 19107, 2003(E).

Juceviciene, V., Towards a Single Information Space for the Environment in Europe, *INSPIRE 2008*, Maribor, Slovenia, 23-25 June 2008.

Ledoux, H., FieldGML: An Alternative Representation for Fields, *annex of ISO/CD 19156, Geographic information — Observations and measurements*, ISO/TC 211 N 2616, 2008.

Peuquet, D.J. A Conceptual Framework and Comparison of Spatial Data Models. *Cartographica*, 21(4), pp. 66-113, 1984.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Towards a Single Information Space for Environmental Management through Self-Configuration of Distributed Information Processing Systems

G. Pavlin, N. J. E. Wijngaards, C.H.N. Nieuwenhuis

Thales Research and Technology Netherlands, DECIS Lab, Delft, the Netherlands
(Gregor.pavlin@icis.decis.nl)

Abstract: This paper illustrates objectives of the DIADEM project which focuses on a novel combination of advanced technologies which facilitate collaborative information processing in environmental management applications. The emphasis is on the principles and tools which facilitate creation of a single information space in advanced systems of systems through a systematic integration of heterogeneous processes. In particular, we illustrate the main principles of the DIADEM Process Integration framework, which supports collaborative processing based on a combination of automated reasoning processes and cognitive capabilities of multiple human experts, each contributing specific expertise and processing resources.

Keywords: *Single Information Space; Self-Configuration; Service Oriented Approach;*

1. INTRODUCTION

The presented paper illustrates the research and development objectives of the DIADEM project (FP7, call ICT-2007.6.3). The project focuses on a novel combination of advanced technologies which facilitate collaborative information processing in environmental management applications. In general, prevention and mitigation of environmental conditions with potentially adverse effects on the population and eco-systems requires (i) identification of critical situations, (ii) impact assessment which takes into account possible evolution of physical processes, (iii) planning and evaluation of countermeasures and (iv) decision making. This can be achieved only through adequate processing of large quantities of very heterogeneous information based on rich expertise about different aspects of the physical world. In order to illustrate the challenges and solutions in such settings we will use a running example. We assume a malfunction at a chemical plant located in a heavily industrialized and densely populated area; a leaking chemical ignites which in turn produces toxic fumes that pose a serious threat for the residents. The situation requires coordinated actions from many operational teams (e.g. fire fighters, police, health authorities, etc.) as well as experts. The ultimate goal is mitigation of the impact the incident could have on health and economy in the affected area. The responsible decision makers have a choice of different approaches: (i) Preventive evacuation of thousands of residents, (ii) keeping the residents indoors or moving them to the shelter and (iii) no action. The presented decision making problem is not trivial. Namely, evacuation can be very expensive and even dangerous; many people have to be moved, roads in and around the area have to be blocked, factories and businesses might have to stop operating, resources for a swift evacuation might be difficult to obtain, first aid services are not available in the critical zone, larger groups of people might panic/revolt, frequent exaggerated reactions can result in insensitivity of the population, etc. Keeping the residents indoors might not be a viable solution if the incident lasts too long, since the toxic fumes could penetrate the houses and late evacuation might be hindered by (higher) concentration of toxic fumes. No action could result in a critical exposure of many people and severe health problems or even loss of lives.

In order to be able to make adequate decisions the decision makers must estimate the actual situation, potential impact on the health, economical consequences, risks associated with the possible actions, required and available resources, etc. Such estimation requires acquisition and processing of large quantities of information. This in turn requires adequate communication and sensory infrastructure as well as reasoning about the relevant aspects of the domain which exceeds cognitive capabilities of a single human. A human expert simply does not have the knowledge over all the relevant mechanisms in the domain and cannot process huge amounts of the available information. On the other hand, full automation of decision making processes in such settings is not feasible since the creation of the required domain models as well as inference are intractable problems. Namely automated inference processes involve many variables and relations with accompanying representation and specific inference mechanisms.

In such settings the solutions must support collaborative processing based on a combination of automated reasoning processes and cognitive capabilities of multiple human experts, each contributing specific expertise and processing resources.

Key to effective combination of human-based expertise and automated reasoning processes is a single information space which allows that each piece of the relevant information is adequately considered in the final decision. The main elements of a single information space are:

- Standardized data formats that facilitate sharing of heterogeneous information.

- Filtering services which provide stakeholders in a decision making process with the information they can process in the context of their role. In principle, filtering services must (i) transform very heterogeneous data/observations to more abstract information types, i.e. extract high-level information through interpretation of heterogeneous cues, and (ii) route the interpretation results to the consumers that can make use of the extracted information.

While a lot of progress has been made with the standardization of data formats and sharing protocols (see INSPIRE, OASIS, ORCHESTRA, MEDSI), currently creation of adequate filtering services remains a major challenge. In this paper we focus on the latter.

In principle, environmental management can be viewed as a special form of crisis management. Consequently, state of the art environmental management is in general plagued by the problems observed in other types of crisis management. For example, the state of the art approaches to crisis management often cannot cope with the complexity of the problems within the typical temporal constraints (see van Santen et al.; [2007] Jennex, [2007]). Beside different data formats, communication protocols, etc., it is often not known who has the relevant expertise or how that person can be reached in a given situation. In other words, an adequate system of experts cannot be established on time and a lot of useful expertise and information is not exploited, which often results in inadequate solutions. On the other hand, full automation of the information processing is often intractable. Namely, due to the domain complexity construction of adequate domain representations (i.e. models) can be intractable and automated inference can be computationally very expensive.

The main objectives of the DIADEM project are the principles and software components which support seamless and systematic combination of human expertise and automated inference processes in complex environmental management domains. This will be achieved through a novel combination of theoretically and technically sound tools and methods for (i) integration of human experts and automated inference in meaningful processing workflows, (ii) automated detection and monitoring of toxic gases and (iii) structuring of human based reasoning about complex domains.

In this paper we illustrate the basic principles facilitating the creation of a single information space for environmental management. In particular, we expose characteristics of the domain and show how modular, service oriented architectures can facilitate the creation of a single information space. In addition, we provide an overview of the technologies that will be incorporated into advanced information processing systems with the help of the presented integration principles.

2. CHARACTERIZATION OF THE DOMAINS

Often effective organizations in crisis management can be viewed as a *Professional bureaucracy* (see van Aart et al. [2004]). Such organizations consist of experts whose services (i.e. expertise) are standardized and known, coordination mechanisms are decentralized while the local processing (i.e. reasoning) is not standardized. Moreover, in such organizations experts of various types, humans and artificial agents, can be viewed as processing nodes in a *workflow* which maps simple sensor measurements and reports from humans to higher level information delivered to decision makers; i.e. each expert implements a particular transformation between different types of information based on appropriate inference processes and domain models (i.e. expertise). Such a workflow consisting of cascaded processes is often present or at least desired in complex crisis management processes. We illustrate this by expanding our running example with a collaboration of experts captured by Figure 1. We assume that the factory staff (FS) at the incident have an overview of the current state of the damaged system; FS can estimate the quantities of the escaping chemical and its type. This information can be used by a chemical expert at the incident location (CE1) to estimate the type and quantity of toxic fumes resulting from the fire. By knowing the location of the fire, the meteorological conditions as well as the quantity and type of the produced fumes, chemical expert (CE2) can (i) estimate the zones in which the concentration of the toxic gases have exceeded critical levels and (ii) identify areas which are likely to be critical after a certain time. She makes use of the domain knowledge about the physical properties of the gases and their propagation mechanisms. In addition, the CE2 guides fire fighter teams (MT) which can measure gas concentrations at specific locations in order to provide feedback for a more accurate estimation of the critical area. A map showing the critical area is supplied to a health expert (HE) who uses the information on population obtained from the municipality to estimate the impact of the toxic fumes on the humans in case of exposure. Finally, the estimated impact on the population is supplied to decision makers, who choose between no action, evacuation and sheltering. This decision also considers estimated time and costs in case of the evacuation from the danger zone as well as the estimated costs and duration of the preventive evacuation. The former estimate is provided by the fire brigades representatives while the later estimate is supplied by the police department. In other words, in such a system each expert can be viewed as module providing predefined services which in turn require services from other experts. Thus, the situation analysis in the presented example can be viewed as a workflow between different, weakly coupled *processing services*, each specialized for specific aspects of the domain (e.g. types and quantities of toxic fumes produced by burning the chemical, propagation of fumes, measurements with chemical probes, etc.). Moreover, a processing service can be provided by a human (e.g. a chemical expert analyzing the extent of the contamination) or by an automated reasoning process (e.g. detection of gases based on automatic fusion of sensor data). Note that, for the sake of clarity, the used example is a significant abstraction of real crisis management processes.

Moreover, each incident is a unique combination of events, requiring a specific workflow consisting of a particular combination of processing nodes. Because sequences of events in incidents are unpredictable, it is impossible to specify an adequate workflow a priori. For example, given the wind direction experts for the evacuation of hospitals and schools might be needed. If the toxic gas is blown to the open sea, instead, no evacuation is required. Thus, in the latter case experts for the evacuation are not needed. However, a crisis can be viewed as an arbitrary combination of *known* types of events/processes, each understood by a human expert or modeled by an artificial agent. For example, the way chemicals burn and react, effects of exposure to toxic fumes, evacuation approaches in hospitals and schools, etc. are independent of the location and time. Therefore, we can obtain general knowledge about such processes/phenomena which can be used for the analysis in any situation involving such phenomena. In other words, mapping between experts and event types can be made a priori, which is the case in the domains we are focusing on (see for example Regionaal Beheersplan Crisisbeheersing [2004]).

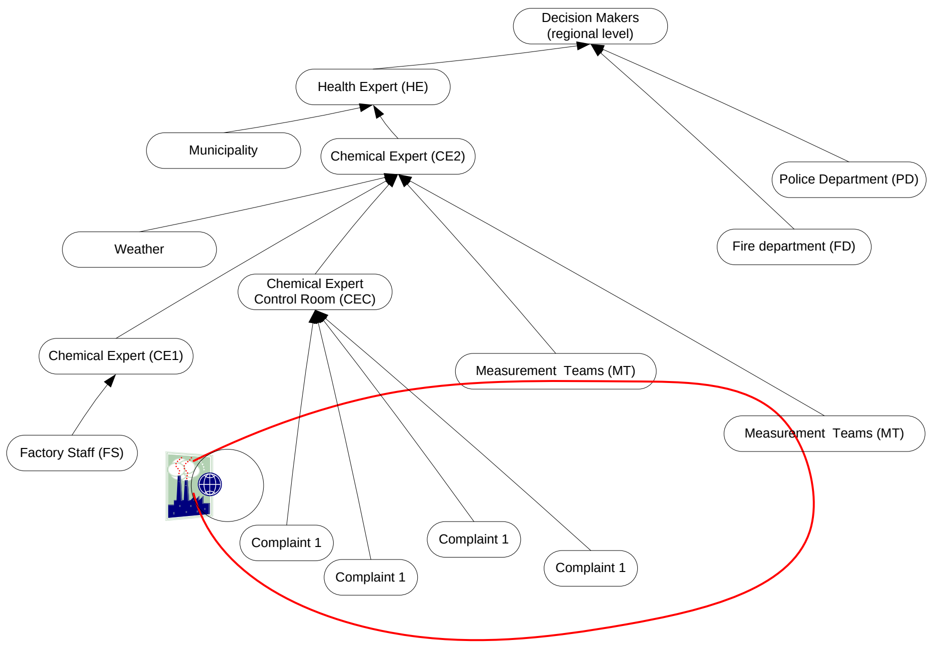


Figure 1: A workflow in a decision making process. Arrows denote information flow between different experts, each processing relevant information of different types. The red region denotes the initial estimate of the area where concentration is likely to be critical.

3. DIADEM PROCESS INTEGRATION FRAMEWORK

The DIADEM process integration framework is a generic service-oriented approach (SOA) which supports efficient combination of very heterogeneous services (i.e. processing capabilities) provided by different experts and automated reasoning processes. A human expert or an automated inference process is represented in the system by a functional module which supports standardized collaboration protocols and, at the same time, allows incorporation of arbitrary reasoning approaches. Thus the modules provide a uniform communication/collaboration infrastructure allowing seamless combination of heterogeneous information processing approaches implemented through human experts or AI techniques. Modules can autonomously form filtering workflows in which heterogeneous processes support collaborative situation assessment, prediction and evaluation of alternatives.

3.1 Processing Workflows

A basic workflow element in the DIADEM framework is a *local process*. Each local process corresponds to a specific transformation between different types of information. For example, a chemical expert CE2 "runs" the local process estimating the hazardous zones, whose outcome (i.e. provided service) is a map of areas which might be dangerous due to high concentrations. In order to accomplish this local processing, the CE2 needs information about the leak, type and quantities of toxic fumes, wind speed and direction, etc. These are the *required inputs* for the local process, which are supplied to CE2 by other modules. Moreover, in our socio-technical view on information processing and decision making systems of systems, it makes sense to abstract from human or machine instances. In the following discussion the term *local process* refers to a reasoning process provided either by a human expert or an automated system implemented by a software agent.

In the DIADEM framework a *local process* is represented by a data structure that describes the provided service, i.e. output of the local process, as well as required inputs (see figure 2.a). The output and inputs are specified either by the human expert providing the processing service or by the designer of the models and algorithms supporting automated reasoning. In case of automated reasoning, the local process description includes also adequate domain models and algorithms.

Since the provided and required services of local processes are explicitly described, meaningful workflows integrating heterogeneous local processes can autonomously be formed on the fly through service discovery and collaboration protocols; each local process that is triggered initiates a search for (i) the providers of the required inputs and (ii) the consumers of the outputs of the local process. The service discovery process makes use of yellow pages, a repository relating service providers and service types. Note that the possible relations between the various services, i.e. types of input and output information for each local process, are encoded in data structures stored by processing modules (see figure 2.a). In other words, no centralized knowledge of the relations between services is required to form workflows supporting globally correct processing. Figure 2.b shows an example of a workflow in which local processes provided by the experts CE1, CE2, MT, HE, W and Municipality (see Figure 1) gradually map heterogeneous observations and sensor data to a high-level estimate of the impact on the health of the population.

With the help of automatic configuration of adequate workflows, the DIADEM framework supports integration of relevant experts into the distributed reasoning processes at runtime, as situation unfolds. In other words, we can cope with unpredictable combinations of known types of events in crisis situation (see Section 2).

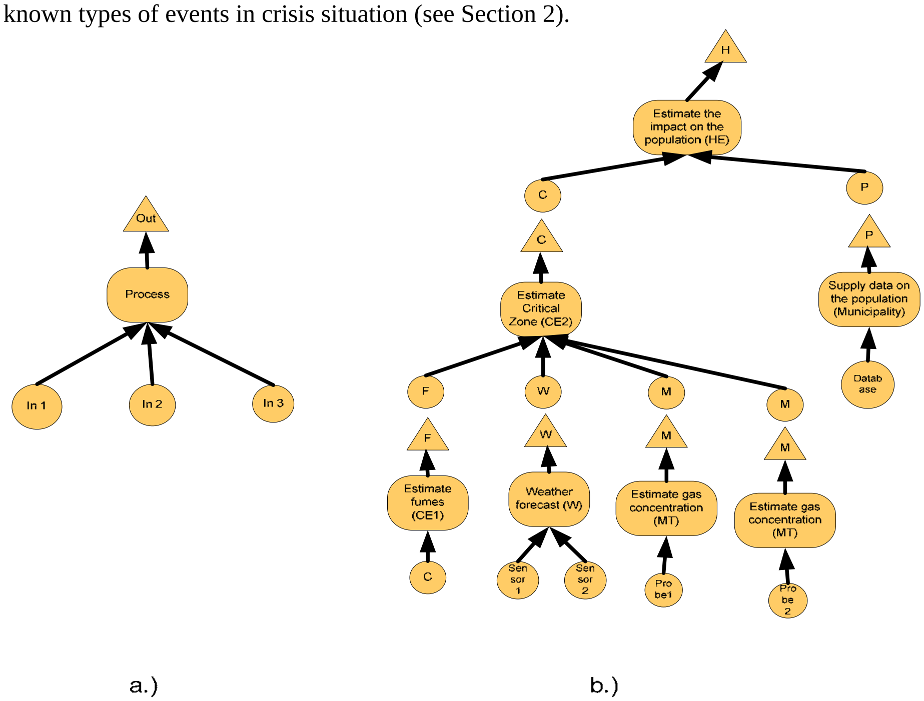


Figure 2: a.) A local process. b.) A system of collaborating local processes forming a workflow specialized for the estimation of the impact of toxic fumes.

3.2 Basic Module Design

Each local process (human or machine-based) is encapsulated by a module which is implemented through a software agent. In the DIADEM framework agents provide a uniform interface between different local processes involved in collaborative information processing workflows. A key feature of the DIADEM agents is asynchronous, data-driven processing in complex workflows. This is achieved through a combination of weakly coupled processes. Each module consists of at least two basic processes implemented through asynchronous threads communicating via a local blackboard (see Figure 3). The *Communication Engine* is a thread which provides inter-module communication, collaboration and negotiation capabilities. *Communication Engine*s in different agents establish workflows between local processes in different agents by executing service discovery and negotiation. Negotiation is based on the *Contract Net Protocol,* Smith [1980]. The *Processing Engine*, on the other hand, is a thread which encapsulates arbitrary automated or human based inference. The *Processing Engine*, implements transformation between different types of information based on interpretation of complex cues. Moreover, each *Processing Engine* can keep track of one or more local processes simultaneously. The *Communication Engine* supplies the *Processing Engine* with inputs and communicates the results of local inference process to other agents.

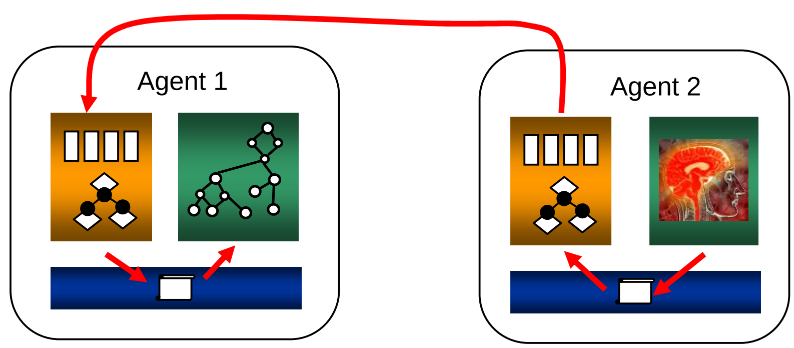


Figure 3: Interaction between agents providing heterogeneous processing services. Both agents use identical communication engine (brown rectangle). However, agent 1 encapsulates automated processing while agent 2 introduces human-based processing.

The *Communication* and *Processing engines* must be able to execute simultaneously. Namely, the reasoning can be computationally expensive which requires certain execution time. At the same time, however, a module should be able to negotiate about possible collaborations with other modules, asynchronously collect outputs from other collaborating modules, etc. Inputs from the local processes in other agents are received by the *Communication Engine* which puts new information to the internal blackboard. This, in turn, triggers callbacks in the *Processing Engine* which supplies the input to the local process that can map the newly obtained information to a particular output (see Figure 3). Upon receipt of a new piece of information, the appropriate process is activated.

Note that, irrespective of the type of the local processes, the Communication Engine and the Processing Engine in each agent use the same mechanisms for the creation and maintenance of correct workflows between local processes in different agents.

The uniformity of configuration mechanisms can be used for a seamless integration of human experts into complex processing workflows. A DIADEM agent representing an expert has a very simple *Processing Engine* which delivers the required inputs to a human expert via a suitable GUI. The expert's conclusions, i.e. his/her service, are also formulated with the help of the GUI and routed to interested agents via the *Communication Engine*. Thus, agents in such a case merely provide automated routing of information between the experts and automated creation of connections between the relevant experts.

Moreover, several types of local processes can be supported by a single DIADEM agent simultaneously. Each process type corresponds to a service an agent can provide. The services are specified by the designers of automated processing modules or human experts who use the agent as their gateway to the DIADEM framework. Each agent automatically registers the specified services as well as required inputs (i.e. services it needs from other agents) at the yellow pages.

3.3 Collaborative, User-Centered System Construction

Efficient construction and maintenance of a single information space is a key to the acceptance of the proposed technology. An important feature of the DIADEM framework is that a new expert can participate in the DIADEM framework by mere specification of the provided services (i.e. the outputs) and the inputs required for the delivery of each service. The agent representing a human expert or automated local processes automatically registers the specified services at the yellow pages. The agents automatically integrate local processes/expertise into the DIADEM workflows through service discovery and negotiation. *Communication Engines* in combination with *yellow pages* provide the infrastructure which supports the service discovery and negotiation. In this approach correct specification of the services is critical; all agents supplying or using the same service must use identical service description. In other words, the services have to be described in a standardized way. However, in the targeted domains services are provided by many stakeholders from very different organizations whose capabilities evolve with the time. Consequently, it is very difficult to specify a complete set of services in advance. In other words, traditional approaches based on a rigorous centralized ontology capturing all relevant services defined prior to putting the system into operation is not practical; we simply do not know which relevant services will be available in the future. Instead, the services should be specified by the users of the system as they upgrade their services. Given the targeted domain, it is likely that the system is enhanced gradually by new experts and automated processing components.

In order to facilitate such an evolutionary, life-long upgrading of the service definitions in the DIADEM-based information space, we introduce the *OntoWizzard* tool, which facilitates collaborative construction of consistent service ontologies. The *OntoWizzard* approach makes use of a central service ontology which is dynamically updated with new service definitions supplied by different service providers. By using the *OntoWizzard*, a designer can describe a new service by using a set of predefined types of concepts; i.e. a service description is a data structure consisting of an arbitrary combination of attributes of limited types. In order to reduce the likelihood of multiple definitions of the same service, the *OntoWizzard* provides a matching service; each designer specifying a new service has to provide a list of keywords as well as a free text describing the service of interest. The *OntoWizzard* uses these informal descriptions to retrieve from the central ontology a list of services that might be identical to the service the designer/expert is trying to specify. The designer can inspect retrieved human-readable descriptions of a limited set of services that are already in the service ontology. If a suitable service exists, the designer adopts the existing description from the central ontology. Otherwise, a new service description is added to the central ontology. In this way the descriptions of services are “synchronized” between the suppliers and consumers of the services.

In other words, the *OntoWizzard* tool facilitates consistent addition of arbitrary service descriptions to the central service ontology, while the format of service descriptions is limited. Note that the configuration of workflows at runtime does not rely on the central ontology; it is used only at design time to facilitate consistent services specification during the configuration of the processing modules.

4. INTEGRATION OF AUTOMATED PROCESSING SERVICES

While fully automated environmental management is not feasible, the overall service quality and performance can be improved significantly through automation of certain processes. In the DIADEM project the current research is focusing on automated gas monitoring and detection. In particular, reliable monitoring at different spatial and temporal scales will be achieved by exploiting the Kernel extrapolation, Gaussian process models and Kalman filter theory. Robustness will be achieved through data-driven approaches to modeling of gas distribution and advanced sensor planning methods. By using sensory measurements from a particular location learning algorithms will automatically extract parameters for local models. Moreover, extended gas distribution models will be used in sensor planning algorithms supporting optimal distribution of sensors. By embedding the monitoring algorithms in the DIADEM framework, the relevant inputs, such as various sensor measurements from different locations, will be delivered to the automated processes and the outcomes will be routed to the interested experts or other automated processes.

Gas detection, on the other hand, will be based on a combination of robust distributed Bayesian systems and novel approaches to integration of human reports and sensor measurements. In our running example depicted in Figure 1, the detectors will partially automate services provided CEC and CE2. Gas detectors will support reliable early alerting based on the existing sensory and communication infrastructure, such as mobile phone networks, Internet, etc. The detectors will be based on the Distributed Perception Networks framework, a modular approach to distributed Bayesian information fusion and information acquisition introduced by Pavlin et al. [2004, 2007]. This framework provides algorithms and modeling methods which support theoretically rigorous automation of local processes implementing distributed gas detection in the DIADEM framework.

5. EFFICIENT COLLABORATIVE HUMAN-BASED REASONING

Higher level analysis and decision making processes often cannot be automated reliably due to the domain complexity. In addition, in mission critical applications humans must be kept in the loop. Therefore, a significant portion of processing in the DIADE M framework will be carried out by humans. Since coherent human-based processing is key to a single information space, a significant effort will focus on tools and methods which facilitate collaborative human-based problem solving. In order to optimally exploit cognitive capabilities of human experts in complex information processing applications, multi-criteria decision analysis (MCDA) and scenario based reasoning methods will be used and extended to distributed settings. Such tools will support systematic structuring of available inference processes. In addition, advanced approaches to human machine interaction will be used in order to seamlessly integrate human cognition into the overall distributed information processing system; information required by experts will be presented in a systematic way and formulation of conclusions will be facilitated by advanced input modalities. MCDA is often chosen as the basis for decision support systems in environmental crisis management (see, e.g., Levy and Taji [2007]) since MCDA aims at providing transparent and coherent support for the resolution of complex decision situations with conflicting objectives. Typically, a limited number of decision alternatives is assessed and evaluated in several respects (criteria). There are no features inherent in MCDA to account for unexpected future developments of the world and system changes.

Scenario planning has been applied predominantly in strategic planning in management, economics, environmental decision making and other areas (see, e.g. Linneman and Klein [1983]). A prime aim of scenario-based reasoning (SBR) is to facilitate the identification of uncertain and uncontrollable factors that may have an impact on the consequences of decisions without turning these uncertainties into probabilities. Scenarios decompose complexity by presenting several alternative plausible developments of the world. SBR strives for finding a strategy that is robust to a range of scenarios that encompass all plausible and relevant futures.

Integrating scenario-based reasoning (SBR) in MCDA is a useful approach as MCDA and SBR balance their respective weaknesses. In DIADEM, Comes et al. [2009] develop an approach that aims at facilitating the decision making under uncertainty arising from possible future developments of an emergency situation (particularly for chemical incidents), exploring the robustness of the decision and allowing for deeper insights into the decision situation.

6. CONCLUSIONS

Significant research and development efforts in the DIADEM project are dedicated to solving the problems of a single information space in complex environmental management processes. In particular, we introduce the DIADEM Process Integration framework, which supports complex information filtering workflows composed out of heterogeneous processes, each implementing a specific transformation between heterogeneous types of information. Such processing in workflows is a basis for advanced reasoning in complex environmental management problems, such as situation assessment, planning and evaluation of options required for rational decision making. This typically involves large quantities of heterogeneous information whose interpretation cannot be carried out reliably by a single human expert or a fully automated system. Instead, the interpretation must be distributed throughout systems of collaborative processing services which form a workflow; each processing node in the workflow provides an interpretation service whose output is used by one or more other nodes. *Meaningful distribution of processing services forms filtering capabilities which are, beside standardized data formats, indispensable for achieving a single information space in complex domains involving large quantities of heterogeneous information.*

Moreover, the DIADEM approach *supports collaborative processing based on a combination of automated reasoning processes and cognitive capabilities of multiple human experts, each contributing specific expertise and processing resources.* This is indispensable since fully automated information processing is not feasible in the targeted domains.

The DIADEM framework is based on a modular, service oriented architecture which supports collaborative processing through combination of automated reasoning processes and cognitive capabilities of multiple experts, each contributing specific expertise and processing resources. The DIADEM framework supports creation of workflows in which arbitrary services, i.e. local processes, can be embedded. *The framework automatically establishes communication between the local processes which can share particular types of information*. The modularity of the DIADEM approach supports distribution of weakly coupled reasoning processes over networked processing nodes. In addition, modularity facilitates gradual (evolutionary) introduction of automated reasoning processes into various organizations involved in mission critical applications; i.e. a “big-bang” introduction of automation is avoided which improves the acceptance.

In addition, the creation of workflows in the DIADEM framework is based on service discovery. This reduces the system configuration to a specification of services that are provided and required by each local process. Each expert specifies which types of information she can provide and which types of information are needed to accomplish a particular service. Similarly, a designer contributing an automated process specifies the provided outputs and required inputs. Therefore, integral components of the DIADEM framework are also methods and tools which facilitate consistent specification of processing services. Such tools support efficient collaborative construction and maintenance of systems involving many developers.

The basic software components of the DIADEM framework that can integrate human processing capabilities as well as Bayesian automated reasoning into complex workflows have been recently introduced. In addition, a simple prototype of the OntoWizzard tool has been implemented. Currently the research and development activities in the DIADEM consortium are focusing on automation of the reasoning processes, negotiation services, human machine interaction and efficient human-based reasoning in distributed settings. The success of integrating human experts and machine reasoning in self-configuring constellations hinges critically on the ability to structure the human-based reasoning and decision making processes – which form an emergent property of DIADEM like systems.

The initial framework provides the first outlines of employing service-configuration for decision-making information processing patterns on the basis of information in the shared information space. The subsequent tooling needs to be aligned at both a more enabling level in the agents' Communication Engine and as GUI's for human participation. In addition monitoring capabilities are required to understand and, if appropriate, influence, the emerging patterns.

ACKNOWLEDGMENTS

The authors wish to thank the colleagues from the DIADEM consortium (www.ist-diadem.eu) for their contributions to this challenging and rewarding research. The presented work is partly funded by the European Commission, grant PJ_REF: 224318.

REFERENCES

Van Aart C. , Wielinga B. and Schreiber G. (2004): Organizational Building Blocks for Design of Distributed Intelligent System, International Journal of Human-Computer Studies, Elsevier, 2004.

Comes, T., Hiete, M., Wijngaards, N. and Kempen. M.(2009). Integrating Scenario-Based Reasoning into Multi-Criteria Decision Analysis. In: J. Landgren, U. Nulden and B. Van de Walle (Eds), *Proceedings of the 6th International ISCRAM Conference*, in press.

INSPIRE (2007): Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) (http://www.ec-gis.org/inspire/

Jennex E. (2007): Reflections on Strong Angel III: Some Lessons Learned, In proceedings of ISCRAM 2007, Delft, the Netherlands.

Levy, J. K. and Taji, K. (2007) Group decision support for hazards planning and emergency management: A Group Analytic Network Process (GANP) approach, *Mathematical and Computer Modelling*, 46, 7-8, 906-917.

Linneman, R. E. and Klein, H. E. (1983) The use of multiple scenarios by U.S. industrial companies: A comparison study, 1977-1981, *Long Range Planning*, 16, 6, 94-101.

MEDSI: http://www.medsi.org

OASIS: Open Advanced System for disaster and emergency management, http://www.oasis-fp6.org

ORCHESTRA: http://www.eu-orchestra.org/

Papamichail, K. and French, S. (2005) Design and evaluation of an intelligent decision support system for nuclear emergencies, *Decision Support Systems*, 41, 1, 84-111.

Pavlin G., de Oude P. and Hood T. (2007): A Modular Approach to Adaptive Bayesian Information Fusion. In *The 10th International Conference on Information Fusion*, Quebec City, Quebec, Canada .

Pavlin G., Maris M. and Nunnink J. (2004): An Agent-Based Approach to Distributed Data and Information Fusion. In IEEE/WIC/ACM Joint Conference on Intelligent Agent Technology, Beijing.

Regionaal Beheersplan Crisisbeheersing 2005-2008, Regionale Hulpverlening Rotterdam-Rijnmond, 2004.

Smith R. G. (1980): The Contract Net Protocol: High-Level Communication and Control in a Distributed Problem Solver. IEEE Transactions on Computers, Series C-29 (12): 1104—1113.

Van Santen W., Janssen A., van der Heijden E., Wijngaards N. (2007): Scenario-Based Analysis for Decision Making: A Condition for Success in Crisis Management, In proceedings of ISCRAM 2007, Delft, the Netherlands.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Automated mapping of environmental variables from a SEIS or SISE perspective

Edzer Pebesma[a], Gregoire Dubois[b] and Dan Cornford[c]

[a] Institute for Geoinformatics, University of Münster, Germany (edzer.pebesma@uni-muenster.de); [b] Joint Research Centre of the European Commission, Ispra, Italy (gregoire.dubois@jrc.it); [c] Computer Science and NCRG, Aston University, Birmingham B4 7ET, UK (d.cornford@aston.ac.uk).

Abstract
The INTAMAP FP6 project has developed an interoperable framework for real-time automatic mapping of critical environmental variables by extending spatial statistical methods and employing open, web-based, data exchange protocols and visualisation tools. This paper will give an overview of the underlying problem, of the project, and discuss which problems it has solved and which open problems seem to be most relevant to deal with next. The interpolation problem that INTAMAP solves is the generic problem of spatial interpolation of environmental variables without user interaction, based on measurements of e.g. PM_{10}, rainfall or gamma dose rate, at arbitrary locations or over a regular grid covering the area of interest. It deals with problems of varying spatial resolution of measurements, the interpolation of averages over larger areas, and with providing information on the interpolation error to the end-user. In addition, monitoring network optimisation is addressed in a non-automatic context.

KEYWORDS: *Environmental data; Environmental information; In-situ sensors, Spatial interpolation, Geostatistics, OGC, SOA.*

1. INTRODUCTION

Spatial interpolation of in situ sensed variables such as meteorological variables, air quality variables, groundwater quality, or environmental radioactivity is a problem for which no single solution exists. In an experiment where several experts were confronted with interpolating the same data set (EUR 21595, 2005), the approaches differed wildly, and best results were obtained by machine learning techniques as well as geostatistical methods. One of the reasons behind this variety is that one needs to choose a model of spatial variability before one can interpolate, and experts disagree on which models are most useful – a case that is not uncommon whenever modelling is involved.

A lack of generally accepted solutions has led to a situation where interpolation experts with highly domain-specific expertise work in fields such as mining, oil exploration, environmental monitoring, or risk assessment and use highly specialised tools. A side effect is that in several domains where interpolation might be useful it is either not applied because of a lack of expertise, or applied using algorithms so simplistic that it undermines the quality of the results.

Motivated on one hand by the increasing availability of sensor data in near real time, and on the other by the need to take decisions in disaster management frameworks without having time to consult interpolation experts, the INTAMAP FP6 project aims to build an automated interpolation service that should provide useful interpolation without requiring any specialised skills. This should be realized using open standards, and under an open source software license. As interpolation cannot be done without introducing errors, the experts in the project consortium considered the word "useful" to mean that the interpolation comes with meaningful information about the interpolation error to characterise the uncertainty in the result. This information might be in the form of an interpolation standard error or prediction variance, the specification of a full conditional probability distribution, or e.g. define probabilities of exceeding a number of given thresholds. Such error information might be ignored by some, but might help others to optimise decision making in the presence of uncertainty, e.g. weighting the risks and costs of type I and type II errors (false negatives or false positives such as evacuating areas not in danger, or not evacuatiing areas that required evacuation).

The paper is organised as follows. First, the statistical considerations underlying automated mapping will be discussed, and the challenges faced outlined. The technical realization and system architecture will be described. Issues of performance and embedding it in a service oriented / service chained environment will be discussed. Finally, we provide a perspective on how this service might be extended along with ideas for future developments of environmental management systems based on service oriented architectures (SOA).

2. STATISTICAL CONSIDERATIONS

Spatial interpolation basically consists of two steps. First, a model for the spatial variability has to be selected, and its parameters have to be estimated. In geostatistics, models of the form

$$Z(s) = m(s) + e(s) \tag{1}$$

are usually deployed (Cressie, 1993), with $Z(s)$ the measured process at spatial location s, $m(s)$ the spatially varying (or constant) trend component usually modelled as a linear in parameters regression model of the form $m(s) = X(s)\beta$ with $X(s)$ often layers in the GIS (Pebesma, 2006) and β unknown regression coefficients, and $e(s)$ usually a second order stationary residual process. This first step then boils down to the choice of a trend function, a covariance function for the residual process, and the estimation of all parameters involved in both components.

The second step involves, given this model and the observations, the spatial interpolation (prediction, evaluation) of this model for new observation locations s_0

$$\hat{Z}(s_0) = \hat{m}(s) + \hat{e}(s) \tag{2}$$

where s_0 is usually taken over a grid covering the region of interest.

2.1 The emergency case: spatial extremes

The original motivation for INTAMAP came from the monitoring of environmental radioactivity at a European scale. EURDEP, the European radiological data exchange platform (see http://eurdep.jrc.ec.europa.eu/), makes unvalidated radiological monitoring data coming from around 4000 sensors spread over most European countries available in near real-time to decision-makers. The main purpose of this network is motivated by emergency cases, where the exchange of these data between EU member states greatly facilitates the monitoring in near real-time of the spread of a radioactive release over Europe. The first stage of an emergency, with a very localised but significant release, is however one of the most difficult problems to interpolate. Several approaches to this have been compared, and developed, within this project. Early stages of a release, such as tested in the interpolation comparison exercise mentioned before (EUR 21595, 2005), are characterised by many low observations and very few observations with extremely outlying measured values. Interpolating such variables is extremely difficult from a statistical perspective.

The INTAMAP automated interpolation service deals with data containing extreme outliers, and deploys dedicated methods, based on spatial copulas, to form a model for spatial variability and interpolate these data (Pilz et al., 2008; Kazianka and Pilz, 2009).

2.2 Uncertainty

Interpolation requires modelling, and modelling involves approximation. Scientist will rarely claim that an interpolated value equals the true value. Statistical models can help to quantify the interpolation error. To interpret the interpolation results in the right way, this information should be transmitted, along with the maps produced. This can be done in several ways, e.g. providing standard errors, probabilities of exceeding thresholds, or by sampling from the statistical model. When the interpolated map is meant to serve as an input to a next processing stage, e.g. to compute exposure of a population over a certain region, this interpolation error specification is indispensable.

The INTAMAP automated interpolation service addresses the communication of errors associated with the interpolation, either in a simple form (standard error, variance) or in a more complete form, ranging from specification of the full parametric distribution (distribution form, parameters), the approximation of this distribution by a number of statistics (e.g. quantiles, or distribution function values), or by a sample from the multivariate distribution (a Monte Carlo sample). The section on technical implementation will detail how this was done.

2.3 Anisotropy detection

Many environmental variables are subject to anisotropy, meaning that in some direction the degree of spatial continuity, or spatial correlation, is stronger than in others. This phenomenon is e.g. present when point sources diffuse, and one transport direction (e.g. due to wind) dominates, e.g. East-West.

The INTAMAP automatic interpolation service automatically detects anisotropy, tests whether it is significant (Chorti and Hristopoulos, 2008), and if it is, corrects for this anisotropy before further steps are taken (modelling of spatial correlation; spatial interpolation).

2.4 Observations with known errors

All observations on continuous variables are measured with some degree of measurement error. Often, this error is unknown, or believed to be very small according to the specifications of the producer of the sensor used. In other cases however, the error magnitudes are known and considerable in size, e.g. because they result from indirect sensing and elaborate and complicated calibration. An example of this are the atmospheric chemistry measurements from satellites such as OMI.

Interpolation of data with considerable, known measurement error should take these errors into account. In the INTAMAP interpolation service, error characteristics of the observations can be specified and a sequential interpolation method based on Gaussian processes (Ingram et al., 2008) is used to optimally interpolate the spatial field in this case.

2.5 Spatial aggregation: estimating areal averages

Besides the usual interpolation to points (on a grid) in space, one may decide to estimate average (or differently spatially aggregated) values, e.g. for complete grid cells, or for larger areas. This may be convenient when decision making does not take place for points, but rather for areas of some size, typically defined by administrative boundaries. An example of this is evacuation: we don't evacuate points, but rather neighbourhoods, regions, villages, towns, or flood plain sections.

The need to consider spatial aggregation in the interpolation process is that although interpolated values can easily be aggregated by averaging them after interpolation took place, the associated errors or error distributions cannot be obtained this way, but need to be quantified during the interpolation process.

2.6 Monitoring network harmonisation and optimisation

Integrating measurements across EU Member States, or even within Member States, e.g. by applying an interpolation procedure often reveals harmonisation issues: different sensor types or different treatment by sensor operation bodies result in constant or random biases. Several possible bias types have been identified, and procedures have been implemented to estimate their magnitude from monitoring network data for the case where they are not reported (Skøien et al., 2009).

Monitoring network optimisation involves the placement, removal or moving of monitoring network stations. Part of this problem is obviously political, as it involves the monetary costs and benefits of a network station, and deals e.g. with questions concerning which variables a society wants to monitor at all. The scientific contribution to this problem involves the monetary assessment of benefits or losses that addition, removal or moving of stations will result in. Implementing a generic, domain independent solution to this is difficult as monitoring networks usually serve different goals (Müller, 2007).

The INTAMAP automated interpolation service does not automatically correct for statistical biases, because a complete understanding and agreement on the magnitude of such biases should be found before it can be part of an automated analysis system. The software delivered does provide the tools to estimate biases from network data. From a similar argument, network optimisation code has been developed but is not interfaced through a web service, because it will mostly be evaluated off-line, most likely in a non-automated setting where many more constraints play a role than an automated service can ever consider.

3. TECHNICAL REALISATION

3.1 OGC Web Services

Web service standards as agreed upon by e.g. ISO TC211, OGC and INSPIRE are the basis for useful generic services that can be part of SISE or SEIS. INTAMAP delivers an interpolation web processing service schematically shown in Figure 1. It accepts sensor data from a sensor observation service (as an observations & measurements document), and returns the interpolation result e.g. as a GML document or as a web coverage service (Williams et al, 2007). To encode the interpolation error information, UncertML, a markup language for specifying information that is represented probabilistically as a representation

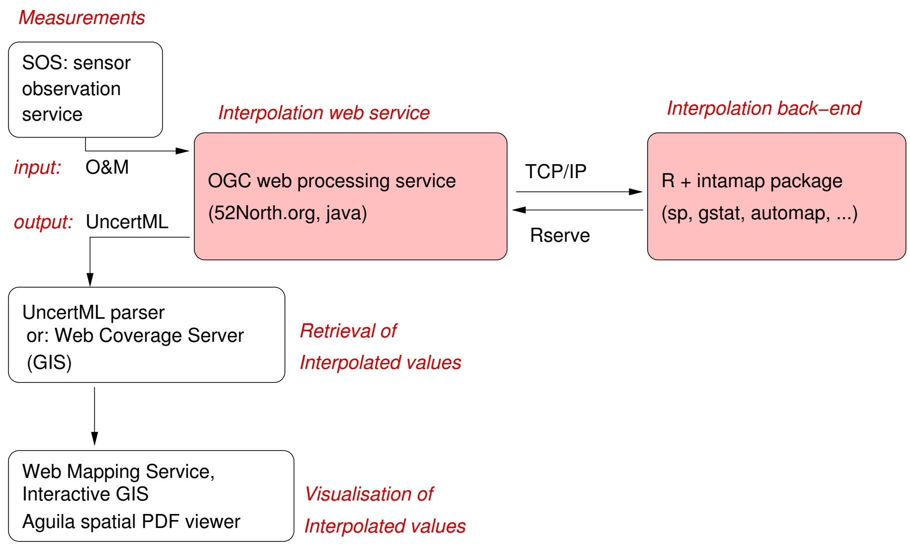


Figure 1: Technical set up of the automatic interpolation service. UncertML stands for uncertainty markup language (see text); O&M stands for observations and measurements, an XML standard for encoding monitoring network data.

of a random variable, has been developed within the project, and proposed as a standard to the OGC (Williams et al., 2008).

3.2 The R back end and interpolation decision tree

The procedure for the statistical analysis of the data are implemented in R, the major open source environment for analysing statistical data. As figure 1 shows, this is not noticeable for the user of the INTAMAP web processing service, as R is run in the back end. Interfacing R from the web processing service by using the TCP/IP protocol (i.e., as a web service, using the Rserve package; Urbanek, 2009) has the advantage that the R process, doing the hard numerical work, may be running on a highly dedicated computing cluster that is not directly connected to the internet. A second advantage of having all statistical routines in the R environment is that it can be re-used independently from the WPS interface, e.g. interactively on a PC, from a python or SOAP interface, or on a mobile device.

The decision tree for choosing an interpolation method automatically is shown in figure 2. In the context of the INTAMAP project, dedicated interpolation methods have been implemented for (i) detecting and correcting for anisotropy, (ii) dealing with extreme value distributions, (iii) dealing with known measurement errors.

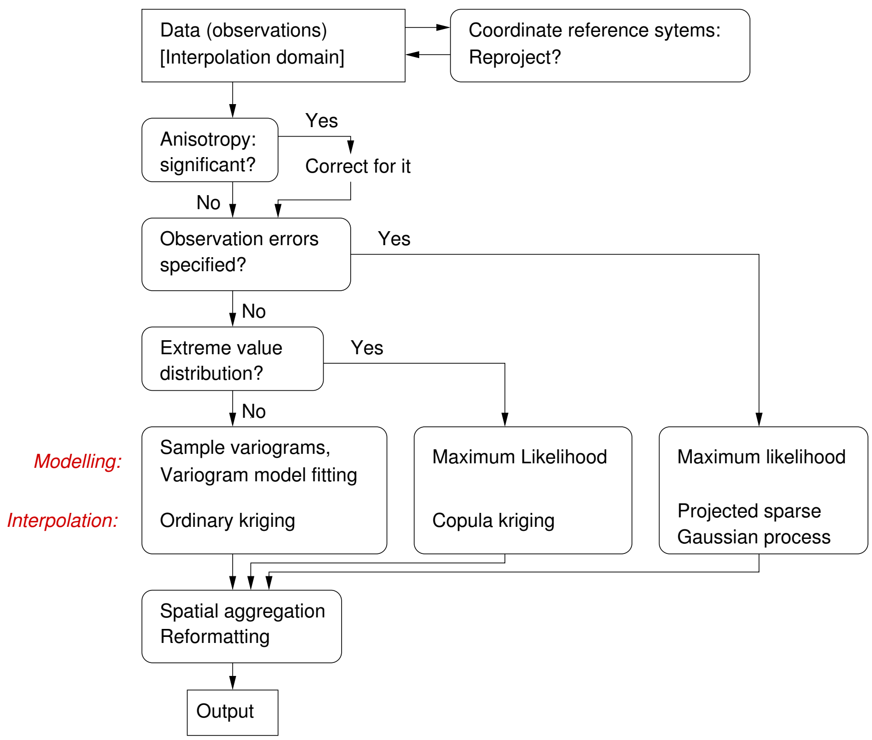


Figure 2: Decision tree for the interpolation method choices in the interpolation process that takes place in R. References in text.

Methods for network harmonisation were also developed, but are not part of the automated interpolation framework, as this should be done before interpolation takes place. The same is true for outlier removal and monitoring network optimisation. With the software developed for INTAMAP, it would be relatively simple to customize the INTAMAP web service and perform these manipulations.

4. OPERATIONAL PERFORMANCE

At the stage of writing this paper, the INTAMAP interpolation service is fully functional, and open for testing. Under this testing framework, the following issues need to be resolved before setting up a robust public service that allows everyone to use it:

- Both maximum likelihood and (global) ordinary kriging need to solve systems of linear equations of size $n \times n$, with n the number of observations. When n becomes large, say over 1000, then this process takes very long. For ordinary kriging this can be solved by reducing the system by default to only address the nearest m observations, with e.g. m in the range of 50.
- When running a web service, it is hard to be certain that the service or server will not at some stage get overloaded when many server requests arrive at the same time.
- Some of the interpolation methods implemented need a considerable amount of time to process, of the order of hours or more; an interpolation request should then specify the (maximum) amount of time available, and the service should be able to discard some of the methods based on that requirement. In case of long-term processes, asynchronous protocols are needed, and these are implemented in the reference WPS used.
- After using the automatic method selection of the INTAMAP service, we envisage that experienced or expert users will want to have more control over the interpolation method chosen. Thus we allow the WPS to accept parameters that are passed to the R process, to control this.
- The observations read by the INTAMAP interpolation service need to be an O&M document (observations and measurements), but not every O&M document will be accepted. This is because O&M accommodates practically every possible observation scenario, including time series data and imagery data – cases that make little sense to send to an interpolation service.
- When used in a controlled environment, e.g. to a restricted domain such as air quality or environmental radioactivity, the R web service can also be constrained to always use the same method, in order to get results that are easier comparable across different interpolation requests.

- Besides interpolated values, the interpolation should return some information about which method was used, what the values of the fitted parameters are, and maybe even some relevant diagnostic plots, such as the variogram and fitted model.

A few observations made here are common to SEIS or SISE. The availability of the web services, the computational time that needs to be accounted for when requests are made in parallel to different environmental web services, the propagation or errors, the tracking and documentation of manipulations in chained service environments are all challenges SEIS and SISE will have to address.

5. DISCUSSION AND OUTLOOK

The automated interpolation web service, the main deliverable of INTAMAP, is an important asset to SEIS or SISE – it takes monitoring data, interpolates to arbitrary points, grids, or averages over polygons, and yields information on the interpolation approximation errors made. It deals with anisotropy, with errors in observations, and with outliers/extreme value distributions. In addition, in a number of application areas (air quality, environmental radioactivity, meteorology) the use of the service will be shown in use cases and demonstrations. The implementation uses open OGC standards and is completely open source. Technology for network optimisation and harmonisation has been developed as well.

The generic interpolation service does just that: automatic interpolation. Clearly, interpolation of real variables with known characteristics would typically not only use measured data, but additional information: for air quality one would like to use remotely sensed data, land use and/or traffic information, for environmental radioactivity it makes sense to use geology and altitude. Although such information is readily available, the appropriate interpolation service would become domain specific (only relevant for a specific variable) and location specific (only useful for a specific region). The generic interpolation service developed here could, however, very well be used as a first major component to build such a specific interpolation service.

In the same thread, phenomena for which near real-time interpolation is relevant are usually dynamic in time, and the interpolation service set up currently ignores time. The step from spatial interpolation to spatio-temporal interpolation is not a trivial one, and again the current development can be used as a first building block for it. One reason not to address time was that in space-time modelling some kind of gradual development of the spatial field over time is usually assumed. In case of unexpected extremes (a nuclear accident), such assumptions may lead to underestimation of the real problems. Further, the behaviour of many variables is subject to transport and diffusion, and involving a transport model would again make the approach domain specific.

For all extension directions: including static GIS information, including dynamic mechanistic models, and including the temporal component, the real challenge lies in developing a method (one or more service) that acknowledges that data are subject to errors, models are subject to errors, and as a consequence spatio-temporal interpolations and model predictions are subject to error as well. These errors should be informative to, and used by, the next level of information uptake, be it modelling or decision making.

Acknowledgements

This work is funded by the European Commission, under the Sixth Framework Programme, by the Contract No. 033811 with the DG INFSO, action Line IST-2005-2.5.12 ICT for Environmental Risk Management. The views expressed herein are those of the authors and are not necessarily those of the European Commission. More information on INTAMAP and UncertML can be found on the internet: http://www.intamap.org/ , http://www.uncertml.org/

References

Chorti, A. and D.T. Hristopulos, 2008. Non-parametric Identification of Anisotropic (Elliptic) Correlations in Spatially Distributed Data Sets, *IEEE Transactions on Signal Processing*, vol. 56(10), pp. 4738-4751, 2008.

Cressie, N.A.C., 1993. Statistics for Spatial Data; Revised Edition. John Wiley & Sons, Inc., New York

EUR 21595 EN, Automatic mapping algorithms for routine and emergency monitoring data. *Report on the Spatial Interpolation Comparison (SIC2004) exercise*. Dubois G. (Ed), European Commission, Office for Official Publications, Luxembourg, 150 p., 2005.

B. Ingram, D. Cornford, and L. Csato, 2008. A projected process kriging algorithm for sensor networks with heterogeneous error characteristics. In: *Geostats 2008 - 8th International Geostatistics Congress*, J. Ortiz & X. Emery (Eds), GECAMIN Ltd., Santiago de Chile, Chile, December 2008.

Kazianka, H. and Pilz, J. (in press), Spatial Interpolation Using Copula-Based Geostatistical Models, In: *geoENV VII - Geostatistics for Environmental Applications*, P. Atkinson et al. eds., Springer, Berlin

Müller, W.G., 2007. Collecting Spatial Data: Optimum Design of Experiments for Random Fields. Springer, New York.

Pebesma, Edzer J., 2006. The Role of External Variables and GIS Databases in Geostatistical Analysis. *Transactions in GIS* Vol. 10 No. 4, 615-632.

Pilz, J., Kazianka, H. and Spöck, G. (2008), Interoperability - Spatial Interpolation and Automated Mapping, In: *Proc. 4th Int. Conference on Information and Communication Technologies in Bio and Earth Sciences HAICTA 2008 (T. Tsiligiridis ed.)*, Agricultural University of Athens, Athens, 110-118

Skøien, J.O., G. Blöschl, E.J. Pebesma (2008) Geostatistics for automatic estimation of environmental variables - some simple solutions. *Georisk*, published online.

Skøien, J.O., O. Baume, E. J. Pebesma, and G. B. M. Heuvelink, 2009. Identifying and removing heterogeneities between monitoring networks, *Environmetrics*, accepted for publication.

Simon Urbanek, 2009. Rserve: Binary R server, R package version 0.4-7. http://www.rosuda.org/Rserve/

Williams, M., Cornford, D., Bastin, L. and B. Ingram. (2008). UncertML: an XML schema for exchanging uncertainty. In: *Proceedings of the GISRUK 2008*, pp. 275-279, April 2008, Manchester, UK.

Williams, M., Cornford, D., Ingram, B., Bastin, L., Beaumont, T., Pebesma, E. and G. Dubois (2007). Supporting interoperable interpolation: the INTAMAP approach. Presented at: *International Symposium on Environmental Software Systems (ISESS 2007)*, Prague, Czech Republic, 22-25 May 2007.

SANY – a European scale project towards shared information

G. Schimak[a] **and D. Havlik**[a]

[a] *Austrian Research Centers GmbH – ARC, 2444 Seibersdorf, Austria*
Gerald.Schimak,Denis.Havlik@arcs.ac.at

Abstract: Information sharing across borders, environmental domains and organisations is one of the strategic goals of the European Commission within Framework Programme 6 and 7. It includes the vision of building a Single Information Space for the Environment in Europe (SISE). This is a challenging task for research and industry as well as for policy.

One of the major topics discussed already at the SANY kick-off meeting (Sensors Anywhere, Integrated Project, FP6; Vienna; Sept.2006), in collaboration with the European Commission (DG INFSO) was about where will SANY go and what is our common vision for the future. Today's keyword of a single information space was more or less created at that kick-off event. Since that time SANY is working on realizing and fostering the vision of a Single European Information Space for the Environment capable to share environmental risk related information. Of course going hand in hand with SISE, SANY addresses also major topics of today's challenges expressed under Global Monitoring for Environment and Security (GMES, now called Kopernikus, announced at the GMES-2008 forum, held in Lille, France; see http://www.gmes.info/ and http://ec.europa.eu/kopernikus/index_en.htm) and clearly takes into account requirements coming from Infrastructure for Spatial Information in the European Community (INSPIRE; http://inspire.jrc.it) as well as from the Shared Environmental Information System (SEIS; http://ec.europa.eu/environment/seis/ index.htm).

This article discusses requirements coming from SISE, SEIS, GMES, GEOSS and map them to SANY service developments (by same use cases) based on OGC Standards. Specifically, this means that we show how these requirements can be implemented through SANY.

Austrian Research Centers (ARC) participates in developing IT infrastructures, architecture and services capable of addressing challenges in the above context as well as in the context of environmental monitoring networks through SANY(www.sany-ip.eu).

Keywords: Sensor Network, Sensor Service Architecture, Infrastructure, SensorSA, Environmental Information; Information Space, Risk Management, Sensor Observation Service, cascading SOS

1. INTRODUCTION

During the past few decades many challenging requirements coming from GMES (Global Monitoring for Environment and Security), GEOSS (Global Earth Observation System of System) INSPIRE (Infrastructure for Spatial Information in the European Community), SEIS (Shared Environmental Information System), SISE (Single Information Space for the Environment) where expressed and influenced heavily ICT developments. In principle all of these programmes, initiatives and systems defined their own information space and requirements. They all build their "cloud" within which they want to collaborate, exchange or share their resources (e.g. information, data or even services). The requirements are very visionary but their realisation is still in a very early stage. We still lake agreed and broadly implemented common architectures, interoperability standards, resource (e.g. information/data/services) management within the Web taking into account practicable and executable IPR issues as well as practical service economy models to make business out of it.

The European Commission (EC) is funding a number of Integrated Projects within the Sixth Framework Programme concerned with exactly above mentioned problems addressing furthermore the improvement and the accessibility of data and services for risk management. ‘Sensors Anywhere’ (SANY, http://sany-ip.eu/) is one important project to mention.

SANY has developed an open distributed information technology architecture and has implemented web services for the accessing and using data emanating, for example, from sensor networks. These developments are based on existing data and service standards proposed by international organizations.

SANY seeks to develop the ideals of the EC directive INSPIRE (http://inspire.jrc.it), which was launched in 2001 and whose implementation began in 2007. Furthermore SANY obeys GEMS, GEOSS (http://www.earthobservations.org) as well as SEIS requirements which you can read in the following chapter.

Thanks to the open nature of the architecture and services being developed within SANY, they can be implemented by any interested party and can be accessed by all potential users. The architecture is based around a service-oriented approach that makes use of Internet-based applications (web services) whose inputs and outputs conform to standards.

The benefit of this philosophy is that it is expected to favour the emergence of an operational market for risk management services in Europe, it eliminates the need to replace or radically alter the hundreds of already operational IT systems in Europe (drastically lowering costs for users), and it allows users and stakeholders to achieve interoperability while using the system most adequate to their needs, budgets, culture etc. (i.e. it has flexibility).

2. INFORMATION SPACE REQUIREMENTS

All large European or world wide Programmes (e.g. GMES, GEOSS) or other Initiatives or Systems (like INSPIRE or SEIS) want to foster the vision of a shared information space. All of them define their shared information space slightly different. But, all of them are more or less converging towards the definition of European Environment Agency’s founded view of Shared Environmental Information System (SEIS; http://ec.europa.eu/environment/seis /index.htm).

SEIS’s objective is the construction of a shared information space for environmental information systems among the Member States [SEC 2008]. This shared information space is used to translate between different existing information systems in order to make them interoperable. Additionally the information exchange is becoming more efficient and current, reliable and relevant data can be provided with virtually no delay, thus enabling decision makers and aid organizations to act sooner and be better informed.

The vision of a Shared European Information Space faces numerous challenges and problems [Ecker, 2008]. Managing the huge amounts of data which are collected and generated throughout Europe is one of them. SEIS addresses this with the requirement that data/information should be managed as closely as possible to the source having in mind that data are merely collected once and multiply shared with others [SEC 2008]. This can be seen as one of the key concepts of a ‘Shared’ Information Space. Making the data available to others is only one aspect; it must also be done in a timely fashion so that decisions can be based on solid data and ideally in real time.

The data must be available and accessible to public authorities anytime and at all administrative levels from the local to the European level (if possible in their natural language) in order to facilitate their legal reporting obligations. But the data must also be available to the general public. The citizens might want to be informed about the state of the environment and to participate in the development and implementation of environmental policies. Abstracting the challenges described before, this results in the following high level requirements [Schimak, 2008] for the SEIS infrastructure implementation:

- Collect data once; Use and share them many times.
- Manage the data at their source level.
- Provide accessibility and availability of data (to public authorities) at any time and at all levels, from local to European, to facilitate reporting obligations as well as to inform citizens about the state of the environment.
- Make information available in relevant national languages

Note: Most of above requirements are referring to the use cases (UC1, UC2, UC4, UC5, UC7) in chapter 3.2.

Just in order to demonstrate how closely other requirements are related we give a short introduction of key architectural requirements specified in the final report for the GMES initial period [GMES FR, 2004]:

“For GMES to become a success, the architecture needs to facilitate the integration of standalone data and information elements. It should allow to the selection and aggregation of information from heterogeneous sources and should provide the capability to translate data and information between the various sources in real time. This applies as much to the incorporation of socio-economic data and information, as well as products derived from the space and in situ observing networks.

Note: Most of this requirements of above paragraph are reflected by the use cases (UC3 and UC4) in chapter 3.2.

GMES must therefore provide a structured framework for data integration and information management, i.e., a European shared information capacity. The following key architectural and user-oriented requirements will therefore drive the implementation of GMES:

- Openness, based on agreed open standards, facilitating seamless communication and interoperability, i.e. the ability of different devices or systems (usually from different vendors) to work together, as well as enabling user service autonomy;
- Federated architecture, enabling systems to grow and evolve;
- Simplicity of architecture (e.g. modularity of components), to break the complexity barrier, systems must be made easier to design, administer and use;
- Self-configuration, programmability, scalability (e.g. to handle various levels of operational load and external conditions);
- Dependability, i.e. the system's resilience to security threats or breakdown;
- User-friendliness of services and interfaces, e.g. in the handling of user request services, access control, workflow management, delivery management, visualisation, data extraction (e.g. “multilinguality”), multiuser sessions, administration;
- Data security, protection of provider and user data against alteration, theft and misuse;
- Quality of service;
- Ubiquity of access, including global reach.”

In a reflection paper [GMES, 2005] on Data Integration and Information Management Capacity, prepared by the EU (DG-INFSO), lists provisionally and non-exhaustively a lot of important functional requirements. The paper suggests that the approach needed to create an efficient data integration and information management component is to use a “system of systems” design. Note: Also requirements coming from GEOSS are following the same interests. All these requirements are reflected in the specification and development of SANY Sensor Service Architecture in Usländer [2008].

The above mentioned system of systems approach (like in GMES or GEOSS) is still regarded today as the best way to go as indicated in the report of Joint Operability Workshop in April 2007 [JOW 2007].

It is further recommended (likewise in GMES, 2005) on Data Integration and Information Management Capacity, the technology of choice to be used to implement the GMES architecture is the Service Oriented Architecture (SOA). An extensive answer to the implications of above addressed requirements can be read in Ecker [2008].

3. SHARING SENSOR DATA

In order to conserve and develop the environment in a favourable way it often depends on detailed knowledge not only about single observable properties, but about ecosystems, its parts, relationships, and dependencies. Building such know-how was and still is often a tedious task, as data acquisition has been expensive and data integration is a labour-intensive task of conversions and transformations, often implying information losses.

SANY IP consortium (http://www.sany-ip.eu) is developing several interesting services that extend the usability of the Open Geospatial Consortium “Sensor Web Enablement” (OGC SWE) architecture. One such service prototype, developed by the Austrian Research Centers, is the “cascading SOS” (SOS-X).

SOS-X is a client to the underlying OGC Sensor Observation service(s) (SOS). It provides alternative access routes to users (or services) interested in accessing, sharing or publishing data. In addition to a simple cascading, SOS-X can re-format, re-organise, and merge data from several sources into a single SOS offering.

3.1 SANY’s underlying concepts

Scaling all these ambitious goals to implementable ones, respectively overcoming the requirements mentioned in chapter 2 this meant for SANY to define and develop an open and standard based Sensor Service Architecture (Sensor SA) plus a set of services (e.g. for data acquisition, data fusion, data sharing, data publishing, etc.) based on available standards.

The Sensor Service Architecture (SensorSA) is a service-oriented architecture (SOA) integrating elements of an Event Driven Architecture (EDA) and having a particular focus on the access, management and processing of information provided by sensors and sensor networks. SensorSA extends the “Reference Model for ORCHESTRA Architecture” (RM-OA) [Usländer, 2007] in the area of in-situ monitoring through:

(1) Inclusion of the OGC “Sensor Web Enablement” service specifications (OGC SWE; see http://www.opengeospatial.org/projects/groups/sensorweb) and

(2) Definitions of the data models and interaction patterns required for in-situ environmental monitoring. In addition, the SensorSA introduces the elements of an Event Driven Architecture (EDA) to the RM-OA and the SensorSA implementation architecture explicitly allows simultaneous use of SOAP, REST-full and standard OGC web services.

Sensor Web Enablement (OGC SWE) is a suite of standardized web-service interfaces and XML schemas that allow live integration of heterogeneous sensor webs into an information infrastructure.

The RM-OA, which has been accepted as a “Best Practice” by Open Geospatial Consortium (OGC) describes an abstract architecture for environmental risk management, a methodology for mapping this abstract architecture onto the technical platform of choice, and a "system of systems" integration concept as way to achieve interoperability on a technical level.

At the level of information sharing and in order to establish interoperability, the RM-OA demands self-description of all resources [Schimak, 2007]. In particular: "all meta-information MUST be provided at least in a form suitable for interpretation by humans"; "syntactic meta-information MUST also be provided in a form suitable for interpretation by machines"; and "providing semantic meta-information in a form suitable for interpretation by machines (e.g. by means of an ontology) is highly encouraged".

The following use cases are demonstrating, according to Havlik [2008] how a cascading SOS is used to collect data/information and to provide them to clients (e.g. humans or other service(s)).

3.2 Use Cases based on cascading SOS

Cascading Sensor Observation Service (SOS-X) is a client to the underlying SOS service(s) and provides alternative access routes to users (or services) interested in accessing data (Figure 1).

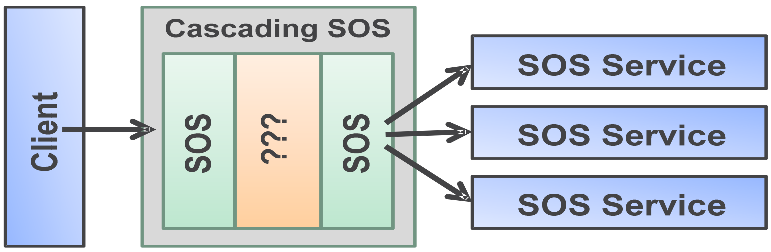


Figure 1. Cascading SOS

In its simplest form, SOS-X provides an alternative route to accessing the data offered by underlying SOS service with no changes to the information model. However, the real power of this service lies in its capability to optimize the data flow, manipulate the meta-information, and to (pre-)process information before re-publishing.

Capabilities of the SOS-X are best understood on the basis of use cases (UC):

UC1: Publishing

In this use case, the cascading SOS is situated in a DMZ and used to publish a subset of internally available data to the outside world (Figure 2a). This mode of operation allows clean separation of the data into "confidential/internal" and "public" part, without the need to replicate the data. In addition, the SOS-X can be instructed to customize the view to the data before re-publishing (UC2, Figure 2b).

UC2: Custom View to Data

In this use case, the cascading SOS is used to provide a single point of access to data from several sources (Figure 2b). The data served by SOS-X itself remains unchanged in this process, but the meta-information and the way data is presented to the users may be altered in the process. This is similar in spirit of ORCHESTRA Translating Feature Access Service (FAS-X) [Anders, 2007], but very different in its implementation. While the FAS-X directly manipulates the XML, the SOS-X translates all incoming data and meta-information into an internal data model.

This architectural decision has one very useful implication: SOS-X can be easily extended to use other sources of information instead of, or in parallel with those available using a Sensor Observation Service interface.

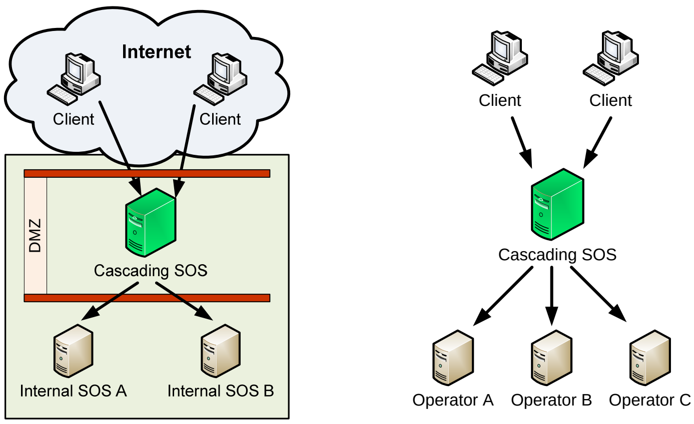


Figure 2. From left to right: (a) UC1 “data publishing”. (b) UC2 “custom view to data”

UC3: Protocol Transducer

Environmental monitoring systems are often used for decades, but the technology often changes very rapidly. As already mentioned in Section 1.2, the SWE specifications were still immature at the time we started the project, and changed in a way that renders the initially used services incompatible with the new developments. This kind of service interfaces incompatibility is a rule rather than exception in the real world.

Cascading SOS can be used to bridge the technology gap between providers and users of information (Figure 2a). As already mentioned in Section 2.2, the SOS-X implementation architecture is modular, and (back-end) clients can be written for various legacy services. In addition, the front-end interface is also implemented in a way that allows easy exchange or even simultaneous access to data over several interfaces.

UC4: SOS Proxy

Information offered by a SOS often contains a large archive of historic data. For example, an air quality monitoring system may contain more than twenty years of archived information. The archive slowly grows over time (e.g. 1 new value per sensor every 30 minutes in case of an air quality monitoring system), and the archived data rarely (if ever) changes.

Archived data is indispensable, e.g. for trend analysis, and training of models or data fusion engines, but fetching it over slow internet connections may be a very time consuming activity. In addition, most SOS server implementations lack support for serving large data sets. The Sensor Observation Service specification in principle allows panning, binary payloads and out of band data delivery. However, neither of these mechanisms is compulsory, and none have been widely used so far.

SOS-X residing on a users LAN could greatly improve the quality of service by pre-fetching and caching the data (Figure 3b). In addition, the SOS-X could implement advanced mechanisms for serving large data sets.

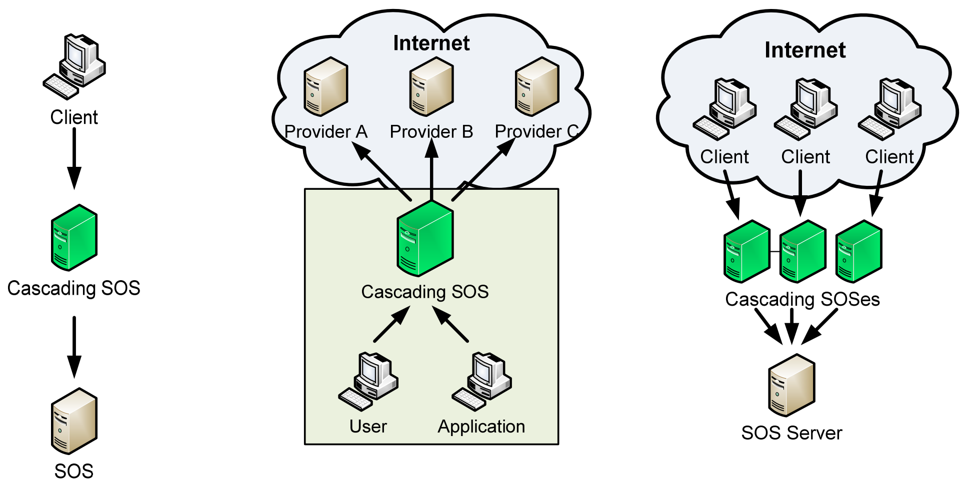


Figure 3. From left to right: (a) UC3 "Protocol Transducer"; (b) UC4 "SOS Proxy"; (c) UC5 "Load Balancing"

UC5: Simple Load Balancing

In emergency situations, the number of requests for information may rise far beyond the average needed server capacity. In addition, most requests concern only a tiny subset of the data. This, and the stateless nature of the SOS service assures that scaling-out is a good answer for emergency overloads.

SOS-X allows a very simple mechanism for scaling out: the original server is moved to the background, and replaced by a group of SOS-X servers. Each of the SOS-X servers is configured to act as exact replica of the original system, and a load balancer assures that the load is evenly distributed over all servers. Thanks to SOS-X caching mechanism, most of the requests can be handled by one of the SOS-X servers, without the need to consult the original SOS service.

UC6: Value Added SOS

This use case is an advanced version of the UC1 and UC2. SOS-X features a built-in mechanism for performing arbitrary algebra operations on time series (Figure 4a). The algebra operations are performed by a time series engine, which is developed in parallel with the SOS-X, and may in the future even be made available as a stand-alone software independent from cascading SOS. Typical (pre-)processing tasks that can be performed by SOS-X include:

- units conversions,
- building of indicators and
- re-sampling of data.

UC7: Sensor Data Store

The "Sensor Data Store" use case can be seen as an advanced version of the "SOS Proxy" (Figure 4b). In use cases 1 to 6, we presume that the cascading SOS does not need to keep a local copy of the original or derived data, except for performance reasons. This implies that:

- the original data will be available on the source server in the future and
- all changes to original data must be reflected in the SOS-X.

These two assumptions are correct in most cases. However, two very important cases exist when these assumptions fail, and the SOS-X needs to manage the data on its own:

First, the SOS-X could be used as a back-end to a decisions support system. Whenever a decision is made, a snapshot of the data the decision was based on needs to be saved for reference purpose. This could be (and often is) done by the decisions support application, but keeping the snapshot of the data in SOS-X would allow more flexibility (e.g. comparing the performance of different algorithms).

Second, some SOS servers may have a limited storage capacity, or even only serve the current sensor value(s). While the information obtained from such servers remains perfectly valid, it may not be available on the original server the next time we need it.

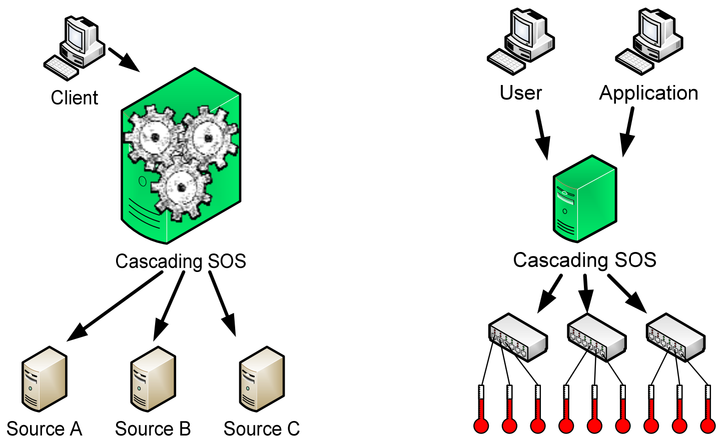


Figure 4. From left to right: (a) UC6 "Value Added SOS". (b) UC7 "Sensor Data Store"

4. CONCLUSION

As you have noticed there are still severe question marks which have to be solved during the next decade, specifically regarding data quality (e.g. when it comes to data fusion), data and service management with in the Web, establishing related policies and last but not least appropriate business models within a new service economy, see also Usländer [2009].

Nevertheless, cascading SOS is a very promising concept, with the potential of becoming a very important infrastructure building block for the in-situ sector of SISE.

The results of the first development cycle are encouraging and no real conceptual problems have been discovered so far. SOS-X is a valuable approach in order to share information, of course in the concert with other (SANY-) services.

However, the software is still in an early stage, and the usability of the final product is highly depending on the performance of the SOS-caching method(s), which cannot be determined before Q3 2009.

In order to improve the visibility of the project, and build communities interested in further development beyond the end of the SANY IP, ARC decided to publish SOS-X under Open Source (GPL license). All information concerning the development status and instructions for downloading and installing the software is available on the SANY-IP web site (http://sany-ip.eu/results/sos_x)

ACKNOWLEDGEMENTS

The authors wish to thank the Information Society and Media DG of the European Commission (DG-INFSO) for co-funding of the SANY and ORCHESTRA integrated

projects within the RTD activities of the Thematic Priority Information Society Technologies.

REFERENCES

Anders, F-S (Ed.), "Translating Feature Access Service (FAS-X)" v0.4, ORCHESTRA – Open Architecture and Spatial Data Infrastructure for Risk Management, Integrated Project, http://www.eu-orchestra.org/publications.shtml#OAspecs , 2007

GMES FR, Global Monitoring for Environment and Security (GMES): Final Report for the GMES Initial Period (2001-2003), http://www.gmes.info/action_plan/index.html, 2004

GMES, GMES reflection paper on Data Integration and Information Management Capacity, DG-INFSO, Draft 6, July 2005

Ecker S. and Schimak G., Fostering the vision of information sharing in Europe in Proceedings of EnviroInfo 2008 - Environmental Informatics and Industrial Ecology; 22th International Conference on Informatics for Environmental Protection, Andreas Möller, Bernd Page, Martin Schreiber (Eds.), ISBN 978-3-8322-7313-2, pp. 204 – 211, 2008

Havlik D., Bleier T. and Schimak G., Sharing Sensor data with SensorSA and cascading Sensor Observation Service in Proceedings of International Workshop Sensing a Changing World 2008, Kooistra, L., and Ligtenberg, A., 2008 (Eds.). CGI report CGI-08-005, Wageningen University and Research Centre, Wageningen, The Netherlands. ISSN 1568-1874, 2008; pp. 80-85

JOW, Joint Operability Workshop Report "Towards a single information space for Environment in Europe", Frascati, 3 April 2007

ORCHESTRA Book, Klopfer, M., Ioannies Kanellopoulos, I. and ORCHESTRA Consortium (Eds.) ORCHESTRA – an open service oriented architecture for risk management; ISBN 979-3-00-024284-7; download able at www.eu-orchestra.org, 2008

Schimak, G., (Ed.) "Rules for ORCHESTRA Application Schemas for Meta-Information", Annex B1 of the Reference Model for the ORCHESTRA Architecture (RM-OA Version 2, Deliverable of the ORCHESTRA Consortium), ORCHESTRA –Open Architecture and Spatial Data Infrastructure for Risk Management, Integrated Project, http://www.eu-orchestra.org/documents.shtml, 2007

Schmak G., et.al. Creating Robust Sensor Networks – Architecture and Infrastructure in Proceedings of iEMSs, ISBN 978-84-7653-074-0, Vol. 3, pp. 1885-1898 2008

SEC Towards a Shared Environmental Information System (SEIS); Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions {SEC(2008) 111, SEC(2008) 112}, 2008

Usländer, Th., (Ed.) Reference Model for the ORCHESTRA Architecture Version 2 (Rev.2.1), an OGC Best Practices Document (07-097) http://portal.opengeospatial.org/files/?artifact_id=23286, 2007

Usländer, Th., (Ed.) "Specification of the Sensor Service Architecture v1", SANY – Sensors Anywhere FP6 Integrated Project deliverable D2.3.2, http://sany-ip.eu/publications/2420, 2008

Usländer Th., Architectural Trends and Viewpoints for the Implementation of the Environmental Information Space, within these Proceedings Towards eEnvironment, 2009

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Prague Environmental Information System - ICT and data sharing in the service of a „City for Life“

P. Štěpánek[a], **V. Kraus**[b], **J. Šolc**[c]

[a] *City of Prague, The City Council Board (petr.stepanek@cityofprague.cz)*
[b] *Prague City Hall, IT Department (vaclav.kraus@cityofprague.cz)*
[c] *SIKS a.s. (jaroslav.solc@siks.cz)*

Abstract: The City of Prague has 20 years of tradition in processing and providing environment related information. The Prague Environmental Information System (IOŽIP) collects and processes environmental data obtained from various sources and thematic projects on the territory of the City. The System provides information outputs to the authorities, experts and the public. A series of annual reports (yearbooks) has been presented on the City's web pages ENVIS together with other information, including the Environment Atlas (maps developed as an interactive GIS application), outputs form Air Quality Modelling, Noise Level Calculation Maps, Vegetation Maps, Environmental Monitoring and other thematic projects. Plans for further development are focused on higher interoperability, data sharing, integration of information presented on the City's portal and specific needs of different end user groups.

Keywords: Prague, Environmental Information System, data collection, sharing and presentation, reporting, GIS, modelling, monitoring, portal.

1. INTRODUCTION

Prague is undoubtedly a fitting place for a discussion about the utilisation of information and communication technologies (ICT) and on environmental data sharing. The conference theme "Towards eEnvironment" could symbolise the long-term efforts of the City in the field of environmental informatics, in harmony with the programme declaration of its councillors aiming at making Prague a "City for Life".

Prague merges aspects of modern metropolitan life with its unique natural environment, coupled with its rich cultural heritage and a tradition of learning. The city's own history and geographical position have allowed it to contribute to the shaping of the history of Europe as a whole. Prague has always had the ambition of being one of the most advanced metropolitan areas. Two decades of democratic development have changed the City in many ways. It has become a successful European region and an attractive tourist destination that protects its monuments as well as precious natural areas. Like any other metropolis, however, Prague faces adverse impacts of the growing automotive traffic, seeks environmentally friendly waste, energy, water and vegetation management; it also strives to regulate the city's expansion into the countryside, optimise the use of mismanaged properties (the so-called "brownfields") and deal with many other issues related to the quality of the City's environment.

The Municipality's planning and decision making regarding issues to concentrate on and allocate available resources for is **based on high-quality objective data**. The data necessary must be collected from a variety of sources, processed using modern technologies and time-tested methodologies and finally shared with end users. The information must be available not only to the City's leadership and professionals but also – in keeping with the policy of transparency - to the general public. It is exactly these tasks that the **Prague Environmental Information System**, abbreviated as IOŽIP in Czech, has been addressing for 20 years.

2. IOŽIP – PRINCIPLES, HISTORY, AND OUTCOMES

The Prague Environmental Information System (IOŽIP) is not a classical agenda system known in the world of ICT in business or government. IOŽIP is a set of controlled activities, a portfolio of projects focused on working with data and outcomes. The system is an example of the Municipality's efforts in the field of eGovernment where ICT tools in combination with organisational changes and skill development provide support to the City Hall's policies and to the information dissemination to the public.

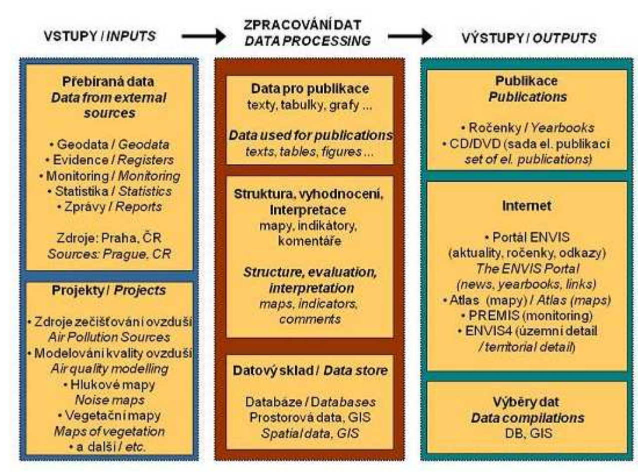


Figure 1. IOZIP scheme – inputs, data processing, outputs

The origin of the system dates back to early 1980s. As time passed, the system underwent numerous organisational as well as technological changes in terms of the tools and instruments used – various databases, geographic information systems, or web technologies and the Internet. Nevertheless, the operation and development of the system have always been ruled by the same principal goals: to collect, process, and provide relevant information to the professional as well as non-professional public in the form of interconnected summary outcomes.

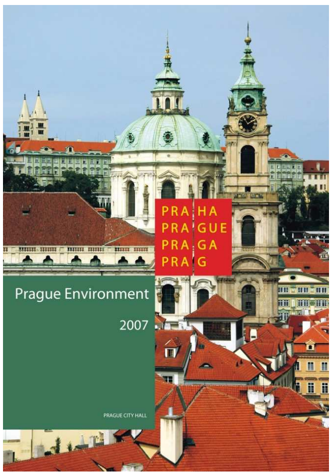


Figure 2. Yearbook Prague Environment 2007

Probably the best known outcome of the system has been the annual publication of the **Prague Environment Yearbook,** first published in 1990 (shortly after the Velvet Revolution). The English version of the Yearbook has been available since 1992. Starting from the late 1990s it has been published in both printed as well as electronic form, available on the City's web site or on CD/DVD, together with other electronic publications. The latest available Yearbook entitled **"Prague Environment 2007"** is the eighteenth part of the series and is also included on **DVD** (enclosed with the conference materials).

In the 1990s, IOŽIP launched **specific projects** using ICT and instruments for presenting, modelling and evaluation of aerial photographs or for transmission and presentation of data from monitoring systems. In the field of Geographic Information Systems (GIS), for instance, the Prague Environment Atlas (hereinafter "the Atlas") has been introduced. Modelling is used in air quality evaluation (project ATEM) and in developing noise level calculation maps. Specific evaluation of multispectral aerial photographs was employed in the assessing of the vegetation health (Vegetation Maps Project). On-line data sharing is utilised in the presentation of data from the network of stations of the Automated Immission Air Quality and Hydrologic Data Monitoring System (project PREMIS). Air quality modelling, noise modelling and landscape and vegetation evaluation in territorial detail have been the goal of last year's ENVIS4 project which brings outcomes for one third of the City so far. The aforementioned projects have their outcomes in the form of specific web applications; basic data are presented in the Yearbooks and in the Atlas.

A thematic information service on the environment called ENVIS was launched in 2005 on the City's web pages (http://envis.praha-mesto.cz). The ENVIS server contains information following the "effect – state – response" scheme, i.e. information on the state of individual environmental components, causes and affecting factors (environment protection instruments), as well as manuals, current news, etc. The web page includes all electronic versions of the Yearbook, links to other web pages, etc.

Over the years, IOŽIP outcomes as well as the aforementioned thematic projects have become an integral part of environmental management in Prague. They have gained recognition and respect not only in the Czech Republic (e.g. in the City Administration, professional institutions, environmental studies researchers, non-governmental organisations, schools and universities, citizens interested in environmental information) but thanks to their English version they are recognized abroad, too.

Foreign cooperation has always been a valued inspiration and opportunity for further progress of the IOŽIP system. The cities of Berlin and Munich played a role as the inspiration behind the creation of the Atlas. Prague's participation in international projects such as CEROI (Cities Environment Report on the Internet), INTERACT, HEAVEN, Air For Europe, CITEAIR, Intelcities, etc., provided valuable experience. Noteworthy is also our capital's involvement with 31 other European cities in the Urban Ecosystems Europe project in 2006 and 2007, which was related to information processing and presentation and evaluated cities on the basis of a set of indicators. Last year Prague, as one of 35 cities, enrolled in the competition of the European Green Capital Award and prepared the entry documentation in the structure required (values of indicators and information on the implemented and planned measures in 11 thematic areas). Experience exchange with other European cities within the network of the EUROCITIES Environment Forum and the Knowledge Society Forum has been used continuously.

3. INFORMATION ON SELECTED PROJECTS AND ACTIVITIES

The following section introduces more detailed information on some of the aforementioned projects that involved the use of technologies and digital data as a crucial part of their solution.

3.1. Prague Environment Atlas Project– a GIS Pioneer on the Web

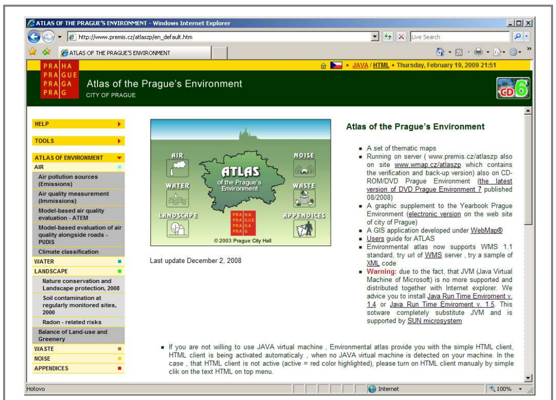


Figure 3. Atlas of the Prague´s Environment

The Prague Environment Atlas Project ("the Atlas") is a geographical web application created over an extensive set of thematic layers, the data base of the GIS-ŽIP (source data in the ESRI Shapefile format). Czech version at: http://www.premis.cz/atlaszp; English version at http://www.premis.cz/atlasen. It contains map views developed by means of the WebMap instrument, arranged into a structure corresponding to the Yearbook. The creation of the Atlas dates back to 1995 when it was first published off-line and printed on paper. Since 1998 it has been available as a web application. The maps are updated depending on data availability from over 20 different sources. Outcomes of other specific projects (ATEM, REZZO, Noise

Maps, Vegetation Maps, etc.), as well as data received from other municipal organisations and the government, are included in the Atlas.

The Atlas currently contains approx. 80 maps created over a foundation of approx. 300 GIS layers, some of which are basic map documents taken over from the data resources of the Prague City Hall. It also enables the "calling" of individual maps from other applications (e.g. from the electronic hypertext yearbook). User tools meet modern standards for geographic web applications (map selection, layer embedding, displaying of information from interlinked databases, on-line editing remarks, etc.); descriptive data to the maps (metainformation) are also available.

3.2. Project ATEM - Air Quality Evaluation by Model Calculations

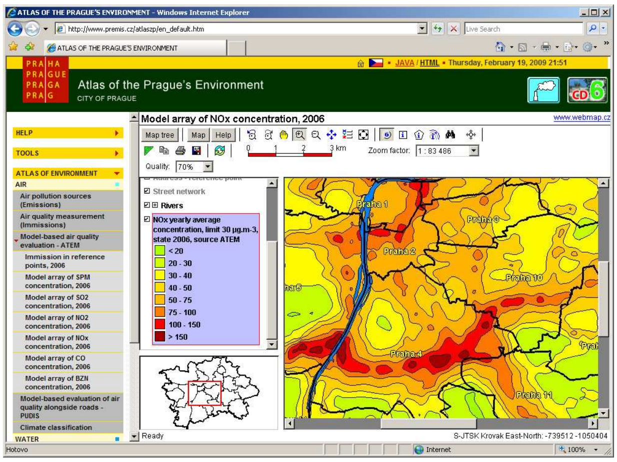


Figure 4. Air quality modelling (the ATEM Project)

The project of the Air Quality Evaluation by Model Calculations was given the name of ATEM after the company that prepared the project for the Municipality in the early 1990s. The project's aim is to determine the immission load on the territory of Prague (annual averages and maximums), level of health risks and shares of pollution sources within the network of calculation points. The first model calculations of the immission conditions for the pollutants of SO_2, NO_x, CO, particulate matter (SPM) and organic compounds (VOC, POP) for the whole territory of Prague were implemented in 1993/94. Since then updating calculations have been carried out every two years (1996, 1998, etc., latest in 2008).

Data on air pollution sources obtained from the REZZO project (Register of Emissions and Air Pollution Sources – i.e. stationary sources - large, medium-sized, and small / households, line sources / transport) on the territory of Prague as well in its surroundings are used as inputs for the calculations. Furthermore, data on dispersion conditions (air flow in ground layers of the atmosphere), digital models of the terrain and territorial data are also employed. In order to evaluate health risks, demographic data are also used. The outcomes include, inter alia, maps depicting the concentration fields of the monitored compounds. The models used were calibrated by comparison with data from measuring stations.

3.3. Noise Level Calculation Maps

The first measurement-based noise maps of automotive traffic in Prague were developed as early as 1976. In the 1990s they were developed on the basis of the combined methodology of measurements and calculations. However, it was only after 1998 that works were launched on a series of Noise Level Calculation Maps of Automotive Traffic, and subsequently also of the Tramway Traffic, as well as their daytime and night-time versions. The calculations were based on the data on traffic intensity on a selected road network, data on road surface quality, speed limits, urban development and vegetation in the vicinity of roads, etc. Demographic data were used for an assessment of the noise disturbance level on the population as part of the follow up works.

In accordance with the European Directive requirements (Directive 2002/49/EC), in 2007 the Strategic Noise Map of Prague Agglomeration was produced and the following year the Action Plan for Noise Reduction in Prague Conurbation was formulated. In 2008, the ENVIS4 project included calculations on "territorial detail", i.e. on a detailed road network.

The calculation models (for both air and noise pollution) have the advantage of lower costs compared to the measurements that would have to be performed in order to evaluate a larger territory. Furthermore, they can be used for environmental impact assessment of planned investments or of environment improving measures. For these purposes the updated calculations for the whole territory of Prague serve as a meaningful starting point and are often used as such.

3.4. Project PREMIS – Prague Environmental Monitoring and Information System

The PREMIS web application (http://www.premis.cz) was created as part of the Prague Environmental Monitoring and Information System in the late 1990s and has been continuously innovated and expanded. The application is focused on the on-line presentation of timely information from the monitoring systems concerning the environment and crisis management.

The main content of the current independent application is up-to-the-hour on-line information on air quality in the Capital City of Prague from the Automated Immission Monitoring system (AIM) informed by the measuring stations of the CHMI and the Public Health Service of the City of Prague, as well as on water levels at the water courses within the territory of the Vltava River Catchment Basin in Prague (including water level predictions for the Czech Republic), basic meteorological information, information from the Radioactive Substance Monitoring (NRPI), and Warning Information.

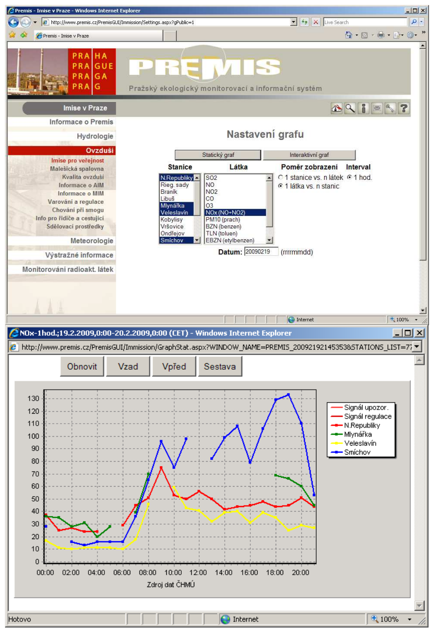


Figure 5. Air quality monitoring data

3.5. ICT and GIS – Helpful in 2002 Flood Control

In the context of ICT, GIS and life in the city there is one more example of what definitely was not a “planned project”. The advantage of the Municipality’s preparedness in the field of technology and digital data on the territory was demonstrated in handling the crisis of 2002, when Prague, as well as a vast territory of Central Europe, was struck by the record-breaking “thousand-year” flood. A crucial role in the relatively successful management of the situation in Prague was undoubtedly played by the timely installation of flood-control barriers, which protected the city’s downtown, as well as the deployment of all components of the Integrated Rescue System, the evacuation of the population, etc. Nevertheless, at that time there was already a map of flood-risk areas (a part of the Land-Use Plan of the Capital City of Prague), developed on the basis of the results of a mathematical flooding model. During the flooding and afterwards the Internet and mobile technologies were at hand as valuable tools for communication with the public (organizing help for those in need, information on transport, on electricity supplies, etc.). Aerial photographs as well as amateur photos were used for the documentation of the deluge, the subsequent assessment of the affected areas and the aid distribution management. On the

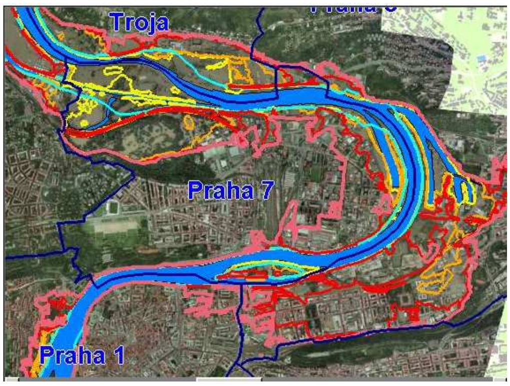


Figure 6. Flood lines, Vltava river 2002

basis of the flood assessment, the municipality proceeded to accelerate the completion of the flood-control system.

4. CONCLUSIONS

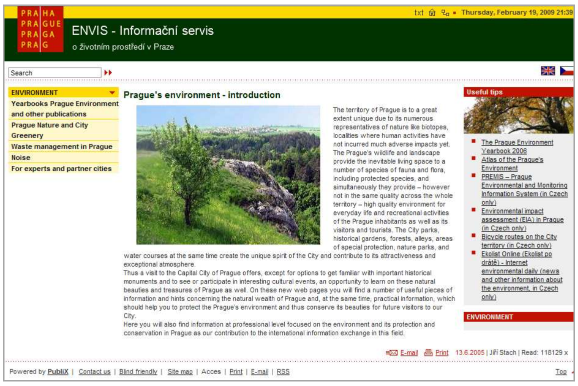


Figure 7. Environmental information servis (ENVIS)

Although the IOŽIP system has already achieved numerous successes and appreciated outcomes, it is clear that there is a need to respond continuously to new stimuli and to the needs of end users as well as development trends. Examples of technological stimuli are, for instance, the expansion of the internet and the increasing computer literacy among the citizens, high-speed networks and mobile technologies, data warehouses and new needs in the organisation of data flows and data management in the city, portal technologies for information presentation and providing the feedback and participation of end users. Requirements for interoperability of information systems and information sharing keep increasing. Methodologies for information processing related to the management of municipal development and environment (indicators) are being standardised. New approaches to education, training, promotion, and increased citizen participation in public administration are emerging.

At present the IOŽIP system's continuity is secured under the conditions of a slightly modified organisational arrangement for dealing with some aspects of municipal informatics. Practical coordination of further system development is in the hands of a task force consisting of representatives of the Prague City Development Authority, the Department of Informatics and the Department of Environmental Protection at the Prague City Hall. An analysis is being made with the objective to assess the current condition and to propose measures for improvement in securing the environmental information flows and administration at the level of input data within the context of building a data warehouse about the city's territory; at the outcome level the objective is to create a new environment for an interconnected and user-friendly presentation of information in the form of the Prague Environment Portal.

Another challenge is to attain a higher level of information sharing within the Municipality (along with Prague City Hall's organisations and local councils) as well as around the Czech Republic by means of providing for interoperability with the systems implemented by the government and its professional institutions and bodies (CENIA, CHMI, NIPH, etc.). Data sharing should meet the European requirements of the SEIS concept (Shared Environmental Information System).

In further development of the IOŽIP system, Prague will continue to strive for a close cooperation with foreign partners, especially with cities and other entities that face similar environmental issues and who endeavour to use information and communication technologies as much as possible in solving them and in informing the public. Prague wishes to become a utile team member on our common journey "Towards eEnvironment".

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http://www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Architectural Viewpoints and Trends for the Implementation of the Environmental Information Space

T.Usländer[a]

[a] *Fraunhofer IITB, Fraunhoferstr. 1, 76131 Karlsruhe, Germany*
(thomas.uslaender@iitb.fraunhofer.de)

Abstract: Interoperability between environmental information systems across organisational, disciplinary and technological borders is one of the key challenges for both developers and stakeholders in the environmental domain. Stakeholders range from public authorities, private companies, consortia of environmental projects, thematic communities up to the citizen. The Single Information Space in Europe for the Environment (SISE) is a vision of the European Commission that promises to enable an easy-to-use but controlled and efficient collaboration between users and providers of environmental information and services. This paper approaches the SISE from an architectural point of view. It positions the SISE implementation problem within a research framework covering both the business and the technological aspects of large-scale environmental information systems. It presents existing and emerging reference models for open geospatial service-oriented architectures (SOA) based upon international standards as the currently established paradigm for distributed environmental applications. Furthermore, it describes selected architectural trends such as Web service paradigms, access control in geospatial SOAs and Semantic Web Services. It is expected that these trends, among others, will influence the discussion about suitable service and information infrastructures for the SISE in the next years. The paper concludes with an outline about the European support action GIGAS that is about to find harmonisation strategies for the different architectural frameworks in strategic European and world-wide initiatives such as INSPIRE, GMES and GEOSS.

Keywords: Environmental information space; Service-oriented Architecture; Reference Model; ORCHESTRA; SANY; GIGAS.

1. MOTIVATION

The Single Information Space in Europe for the Environment (SISE) is a vision of the European Commission that, once having been implemented, shall provide the infrastructure for the efficient, easy-to-use but controlled collaboration between users and providers of environmental information and services. The stakeholders range from public authorities, private companies, consortia of environmental projects up to the citizen. Organisational, disciplinary and technological borders shall be overcome in a rather seamless manner. The plethora of environmental data collected in environmental monitoring programs at several organisational levels shall be made available for further processing, information fusion, visualisation, reporting and decision support, not only for the environmental domain but also for associated thematic domains such as health, security, commerce and transport. Interoperability is a major concern in setting up the SISE [Coene and Gasser, 2007].

Such an ambitious vision needs a basic Information Technology (IT) architectural framework that enables a cost-effective but innovative implementation:

- It shall integrate existing environmental information systems on local, regional, national and international level.
- It shall be secure, dependable and flexible to cope with and integrate emerging technological trends.
- It shall be aligned with large-scale European initiatives such as the Infrastructure for Spatial Information in the European Community (INSPIRE) [EC, 2007], the Global Monitoring for Environment and Security (GMES) and the Shared Environmental Information System (SEIS)" [EC, 2008], but also with world-wide initiatives such as the Global Earth Observation System of Systems (GEOSS).

There is no single way of specifying an SISE architecture, instead, an architecture of such a complexity is typically specified from several viewpoints. The use of viewpoints is derived from the principle of abstraction as the heart of architectural specifications. According to the ISO Reference Model for Open Distributed Systems (RM-ODP) [ISO 1998], a viewpoint is a "form of abstraction achieved using a selected set of architectural concepts and structuring rules, in order to focus on particular concerns within a system."

Firstly, this paper sketches the research framework (section 2) for the SISE implementation. Secondly, reference models that are relevant for the specification of the SISE architecture are presented (section 3). Finally, the paper discusses architectural trends (section 4) and concludes with a summary and outlook (section 5).

2. SISE RESEARCH FRAMEWORK

When designing information systems, there is a dualism between the organizational design to create an effective organizational infrastructure that is derived from the business strategy, and the information system design to create an effective information system infrastructure derived from the IT strategy. Hevner et al [2004] state that on the one hand, "IT are seen as enablers of business strategy and organizational infrastructure", on the other hand "available and emerging IT capabilities are a significant factor in determining the strategies that guide an organization".

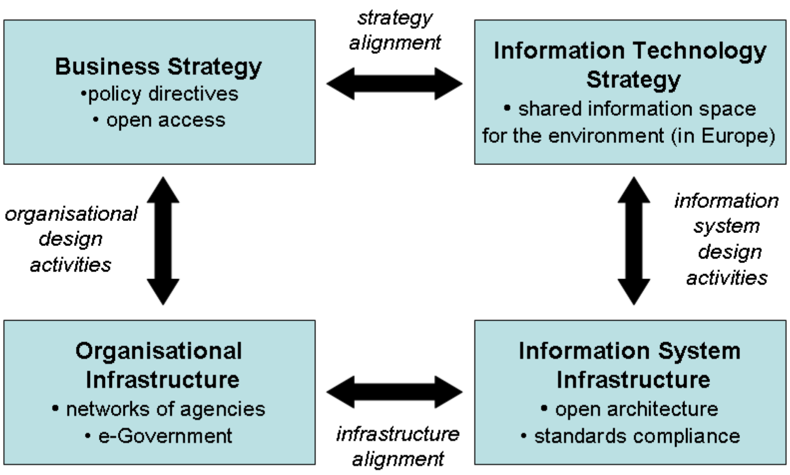


Figure 1. Business and IT Strategies [Henderson and Venkatraman, 1993], adapted to SISE

Based upon the original "strategic alignment model" of [Henderson and Venkatraman, 1993], Figure 1 illustrates the impact of European policies and initiatives upon the strategies and infrastructures of the SISE design activities:

- The Business Strategy for environmental risk management is mainly driven by European and national environmental legislation resulting in policy directives such as INSPIRE, SEIS or the Directive on Public Access to Environmental Information [EC

2003). These strengthen the need to exchange environmental information between the stakeholders and to enable the access to environmental information to the public.

- Translated to the IT domain, the Information Technology Strategy for the SISE is mainly determined by the SISE ambition. The SISE is defined both as a vision and a need towards which all IT research activity in this domain has to be directed.

These strategies determine the design activities on the organisational and the technical side:

- The “design” of the Organisational Infrastructure shall support European environmental legislation. Environmental agencies and ministries, research institutes and private organisations are organised as nodes in networks (e.g. networks of excellence for selected thematic domains in e-Government relationships) with defined levels of cooperation on managerial and technical level.
- An essential requirement for the Information System Infrastructure is an “open architecture” for a service platform which provides seamless access to information, services and applications across organizational, technical, cultural and political borders. “Open” hereby means that service specifications make use of existing standards where appropriate and possible, are published and made freely available to interested vendors and users with a view of widespread adoption.

3. ARCHITECTURAL VIEWPOINTS

The design of information system infrastructures is heavily related to the architectural evolution of environmental information systems (EIS) which have to be integrated. In the last 10-15 years, the EIS design has undergone fundamental changes following the need to correlate environmental information and services across various thematic domains, open it up to a wider spectrum of users (from employees in environmental agencies, over politicians in ministries up to the citizen) and make more sophisticated functions directly available within an EIS, such as environmental simulations or geo-processing capabilities [Usländer, 2008a].

Service-oriented architecture (SOA) is the design paradigm that is currently aimed at large-scale EIS. The major SOA design principles [Erl, 2008], the loose-coupling of functional entities (services) on the basis of an agreed service platform, their autonomy, reusability and composability, fosters its use in geospatial information systems. It enables to share geospatial resources, i.e. data and services with an explicit or implicit geospatial reference, to compose them to higher-level entities and to use them in geospatial applications possibly distributed across organizational and administrative boundaries. This is essential for EIS as natural phenomena are not limited to boundaries drawn by humans.

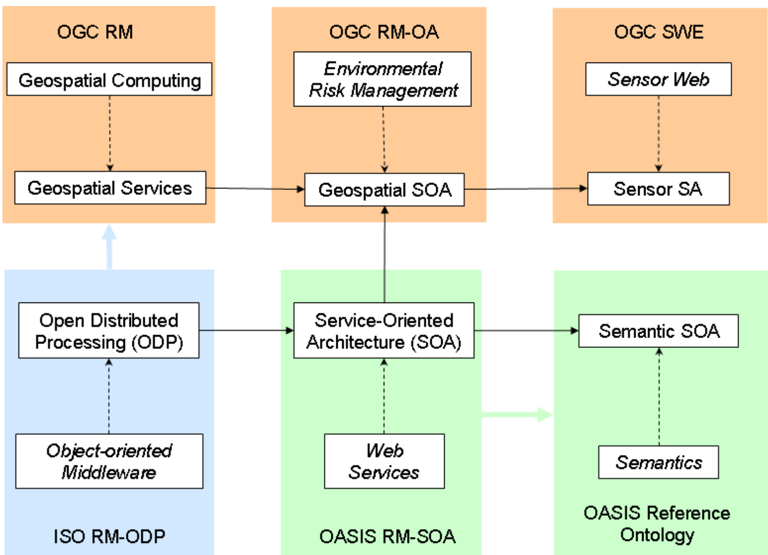


Figure 2. Evolution of Reference Models (RM)

However, as in many other businesses, the transition to service-oriented solutions is still ongoing. SOA requires significant changes in business process design as well as in modelling and solution development as, more than any other objective, SOA is intended to create a stronger alignment between IT and the businesses it supports [High, Krishnan and Sanchez, 2008]. Thus, large-scale EIS result in a system-of-systems architecture that spans multiple organizational, national and technological barriers. A recent example is the world-wide initiative to create GEOSS that aims to integrate existing earth observation systems into a global system that can be applied to various areas of environmental science and management.

For many years, the European Commission has supported research activities to analyse user and system requirements, to derive architectural principles, and to specify and implement generic components of an IT architecture for a future SISE implementation. Various approaches have been launched in order to stimulate a market based on agreed architectures resulting in a series of reference models (Figure 2). They set the conceptual foundation of distributed systems and especially SOAs, with the latest extensions towards geospatial and sensor-related SOAs worked out by the Open Geospatial Consortium (OGC), and, in parallel, extensions towards semantically-enabled SOAs worked out by the Organization for the Advancement of Structured Information Standards (OASIS), see Fensel et al [2008].

As a starting point, inspired by “distributed processing systems based on interacting objects”, ISO has defined a Reference Model for Open Distributed Processing (RM-ODP) [ISO/IEC 1998]. The RM-ODP constitutes a way of thinking about architectural issues in terms of fundamental patterns or organizing principles, and it provides a set of guiding concepts and terminology for building distributed systems in an incremental manner. The major idea of RM-ODP is to structure the documentation of a distributed architecture according to five viewpoints (see the left-hand side of Table 1). The RM-ODP standards have been widely adopted.

The European research project ORCHESTRA [Klopfer and Kannellopoulos, 2008] has applied the RM-ODP approach to the specification of an Open Architecture and Spatial Data Infrastructure for Risk management. The resulting Reference Model for the ORCHESTRA Architecture (RM-OA), accepted as OGC best-practices document, has interpreted the RM-ODP viewpoints as described in table 1 [Usländer, 2007].

Table 1. Interpretation of the RM-ODP Viewpoints for the Design of a Geospatial SOA

RM-ODP Viewpoint	Definition according to ISO/IEC 10746	Interpretation for a geospatial SOA
Enterprise	Concerned with the purpose, scope and policies governing the activities of the specified system within the organization of which it is a part.	Reflects the analysis phase and results in a documentation of the functional, informational and qualitative requirements.
Information	Concerned with the kinds of information handled by the system and constraints on the use and interpretation of that information.	Specifies the modelling approach of all categories of information including their thematic, spatial, temporal characteristics as well as their meta-data.
Computational	Concerned with the functional decomposition of the system into a set of objects that interact at interfaces – enabling system distribution.	(referred to as Service Viewpoint) Specifies the modelling approach for Interface and Service Types including their syntax (signature) and semantics (functional effects).
Technology	Concerned with the choice of technology to support system distribution.	Specifies the technological choices of the service platform, its characteristics and its operational issues.
Engineering	Concerned with the infrastructure required to support system distribution.	Specifies the mapping of the service specifications and information models to the chosen platform and (geospatial) service networks including operational policies.

The RM-OA is built upon two main pillars: a conceptual model and a process model. The conceptual model provides a uniform meta-model including a set of rules how to specify information models, interfaces and services. The process model applies an incremental, iterative approach for the analysis and design phases. Usually, a multi-step breakdown process across several abstraction layers is necessary to analyse the functional, informational and non-functional user requirements and map them to the capabilities of a service platform. In practice, the individual process steps are often interlinked.

The RM-OA distinguishes between an abstract and a concrete service platform. The abstract service platform is specified independently of a given middleware technology. Its assumptions about the capabilities of the underlying concrete service platform are formulated in terms of the abstract SOA concepts of the OASIS Reference Model for Service Oriented Architecture [OASIS, 2006]. The most prominent example of a concrete service platform is the W3C Web Services technology based on the Web Service Description Language (WSDL) and bound to the SOAP protocol.

The European research project SANY (Sensors Anywhere) has extended the RM-OA to a Sensor Service Architecture (SensorSA) [Usländer (ed.), 2009] by the inclusion of sensors, sensor networks and related services based upon the information models and service specifications of the OGC Sensor Web Enablement (SWE) initiative [Simonis, 2008].

4. ARCHITECTURAL TRENDS

As pointed out in the research framework in section 2, there is a dualism in the definition of the IT strategy for the SISE and the technological capabilities for the design of the underlying infrastructure. Some of the emerging and ongoing trends are described below.

4.1 Web Service Paradigms

Looking at the technical foundation for service platforms, there is an ongoing discussion about the basic Web service paradigm to be used. OGC services are typically still used with an http/key-value pair binding although the OGC Technical Committee has decided in 2006 to provide additional WSDL/SOAP bindings for the OGC service interfaces. Furthermore, the mass-market in the Geospatial Web tends towards another paradigm, the RESTful web services [Richardson and Ruby, 2007]. RESTful web services aim at accessing and manipulating uniquely identified resources based on a uniform interface with commonly agreed, well-defined semantics such as the http-protocol of the World Wide Web.

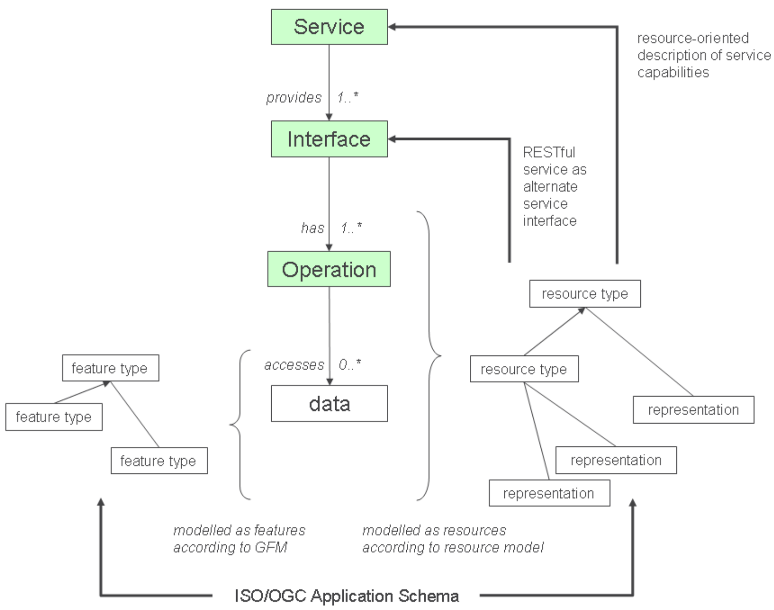


Figure 3. Services, features and resources and possible relationships

The OGC Architecture Working Group is discussing relationships and mapping possibilities between different service paradigms. The author of this paper claims that the conceptual relationships between OGC features, services and resources cannot be described without having a purpose of the resource modelling in mind. Figure 3 illustrates possible relationships derived from the basic assumption of the RM-OA that a service provides one or more interfaces, and each interface consists of one or more operations. Operations access underlying data in a read and write mode. The left-hand side of Figure 3 shows the traditional approach of OGC services accessing underlying data whose structure is modelled as an ISO/OGC application schema. The right-hand side shows a complementary resource-modelling approach. Here, operations are modelled together with their underlying data in form of resources and their representations. There are two possible purposes and applications for this modelling approach:

- The capabilities of a service may be specified in a resource-oriented way. Typically, the resulting resource model mirrors the basic elements of the underlying application schema of a service. As an example let's take the OGC Sensor Observation Service (SOS) whose capabilities may be as resources that reflect the basic SOS concepts such as offerings, features of interest, observation collections or observed properties.
- An extended approach is to provide a RESTful service as an alternate interface based on a resource model on top of a service instance (or by combining several service instances). Such a RESTful service would then provide a selected view (typically just a subset) upon the capabilities and operations of the underlying service.

Resource modelling may provide a modelling bridge between the Information and the Service Viewpoint in a system design. Resource-oriented modelling of services may facilitate the understanding of the functionality of the service to a system designer as the semantics of the operations upon the resources is simple, generic and well-defined. Furthermore, the provision of a resource-oriented view upon a service may facilitate the discovery of the service as the notion of resources may be closer to the "universe of discourse" of the user as it is the case for the signatures of specific services.

4.2 Access Control

Security is a cross-cutting concern in setting up an SOA-based infrastructure for the SISE. The most basic challenge is the protection of resource against unauthorized access as this is the foundation of most other security concepts and adjunct topics such as licensing, digital rights management and protection against malign system interaction. Access control mechanisms ensure that only authorized users (subjects) may access resources using well defined methods that comply with the security policy of the system. Access control solutions for geospatial SOAs are typically realised according to the abstract access control pattern introduced by the eXtensible Access Control Markup Language (XACML) of OASIS [2005], see Figure 4.

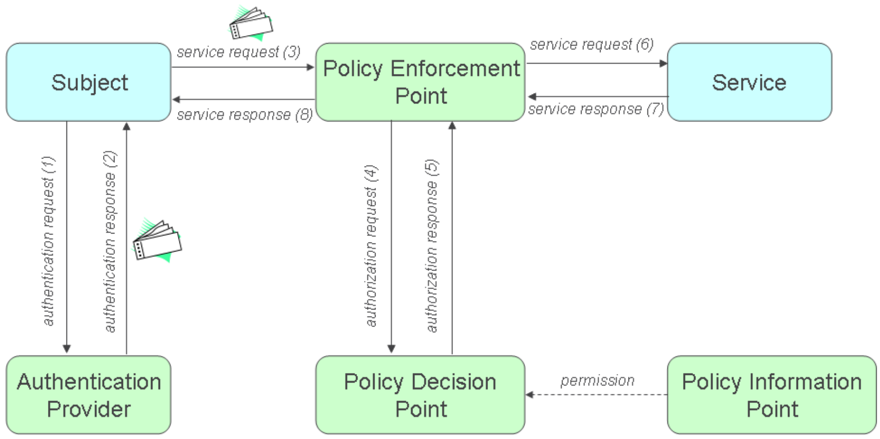


Figure 4. Abstract Access Control Pattern according to XACML [OASIS 2005]

It seizes on general ideas of policies as the basic mechanism to support the governance of SOAs and assumes that the permission "who may access which resource using which action" are recorded in access control policies. The data flow is as follows:

- Before the Subject may issue a service request to a Service, he has to authenticate his identity. He delivers credentials (e.g. password or a key) (1) to an Authentication Provider, and receives a ticket as proof of his successful authentication (2).
- The Policy Enforcement Point receives a service request associated with the ticket (3), but first has to enforce the access control policy (4), (5) before forwarding the service request to the Service (6) and getting back the service response (7), (8).
- The Policy Decision Point responds to an authorization request (4) with an authorization decision (5) based upon permissions of the Policy Information Point.

4.3 Semantic Web Services

Research work on semantic extensions of Web Services has resulted in competing submissions of sophisticated Semantic Web Services frameworks to the W3C as well as corresponding ontologies and semantic execution environments in OASIS [Fensel et al, 2008]. As a first step, the W3C recommends to use "Semantic Annotations for WSDL and XML Schema" (SAWSDL) [Farrell and Lausen, 2007] as a set of extension attributes for the WSDL and XML Schema definition language (Figure 5). The approach is to annotate elements of WSDL documents, in particular interfaces and operations as well as their input, output and fault message structures. This is realised by model references to concepts in semantic models, e.g. ontologies. SAWSDL enables service and data matching on semantic level. However, the interoperability gain has still to be validated in practical use cases.

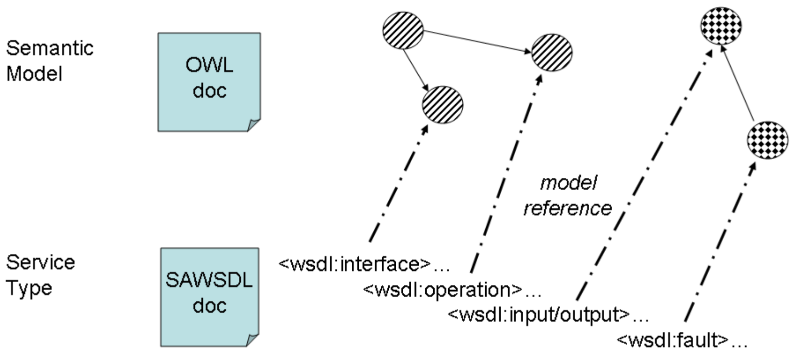


Figure 5. Model References of the Semantic Annotations for WSDL and XML Schema

5. CONCLUSION

In order to integrate the SISE vision into the business strategies of the stakeholders there is a need for a sound architectural basis that, on the one hand, enables interoperable environmental applications today, and, on the other hand, incorporates extension points to leverage emerging and future architectural trends. Key enablers for SISE are open standards-based geospatial information and service platforms. A harmonised IT strategy of the major stakeholders in the various private-public partnerships, initiatives and research programmes is essential to foster interoperability, the ambition of the European project GIGAS [2009]. GIGAS is a two-years coordinated support action that promotes the coherent and interoperable development of the GMES, INSPIRE and GEOSS initiatives through their concerted adoption of standards, protocols, and open architectures. GIGAS is analysing the architectural approaches and solutions for important use cases (e.g. resource discovery) of the above mentioned initiatives as well as major European research projects. GIGAS has adopted the ISO RM-ODP approach to streamline architectural discussions according to viewpoints that focus on selected concerns and different levels of abstraction.

The paper has shown that the technological capabilities are available and may be taken up to make the SISE vision a reality. However, a continuous alignment of business and IT strategies with well defined iteration milestones, as well as harmonised design activities for both information system and organisational infrastructures are indispensable to ensure the return on investments that is expected by the stakeholders.

ACKNOWLEDGEMENTS

This work has been supported by the European Projects ORCHESTRA, SANY and GIGAS co-financed by the Information Society and Media DG of the European Commission.

REFERENCES

Coene, Y. and R. Gasser, Towards a single information space for Environment in Europe, Joint Operability Workshop Report, Frascati, ftp://ftp.cordis.europa.eu/pub/ist/docs/environment/sise-workshop-report_en.pdf, 2007.

EC, Directive 2003/4/EC of the European Parliament and of the Council on public access to environmental information and repealing Council Directive 90/313/EEC, 2003.

EC, Directive 2007/2/EC of the European Parliament and of the Council establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), 2007.

EC, Communication from the Commission to the Council, the European Parliament, The European Economic and Social Committee and the Committee of the Regions - Towards a Shared Environmental Information System (SEIS), COM (2008) 46 final, 2008.

Erl, T., SOA: Principles of Service Design. ISBN 0-13-234482-3. Prentice Hall, 2008.

Farrell, J. and H. Lausen (eds.). Semantic Annotations for WSDL and XML Schema. W3C Proposed Recommendation, http://www.w3.org/TR/sawsdl/, 2007.

Fensel, D., M. Kerrigan and M. Zaremba (eds.), Implementing Semantic Web Services - The SESA Framework. ISBN 978-3-540-77019-0, Springer, 2008.

GEOSS, Global Earth Observation System of Systems, http://www.earthobservations.org

GIGAS, GEOSS, INSPIRE and GMES - an Action in Support, http://www.thegigasforum.eu, 2009.

GMES, Global Monitoring for Environment and Security, http://www.gmes.info.

Henderson, J., and N. Venkatraman. Strategic Alignment: Leveraging Information Technology for Transforming Organizations, IBM Systems Journal (32:1), 1993.

Hevner A.R., S.T. March and J. Park, Design Science in Information System Research. MIS Quarterly Vol. 28 No. 1, pp.75-105, 2004.

High Jr., R., G. Krishnan, and M. Sanchez, Creating and maintaining coherency in loosely coupled systems. IBM Systems Journal, Vol. 47, No 3, pp. 357-375, 2008.

INSPIRE, Infrastructure for Spatial Information in Europe, http://inspire.jrc.ec.europa.eu.

ISO/IEC 10746:1998 (E). Information technology - Open Distributed Processing – Part 1: Reference model, Part 2: Foundations, 1998.

Klopfer, M. and I. Kannellopoulos (eds.), ORCHESTRA – An Open Service Architecture for Risk Management, ISBN 978-3-00-024284-7, 2008.

OASIS, eXtensible Access Control Markup Language (XACML) V2.0, OASIS Standard, http://docs.oasis-open.org/xacml/2.0/access_control-xacml-2.0-core-spec-os.pdf, 2005.

OASIS Reference Model for Service Oriented Architecture 1.0. Committee Specification 1, http://www.oasis-open.org/committees/download.php/19679/soa-rm-cs.pdf, 2006.

Richardson, L., and S. Ruby, RESTful Web Services. O'Reilly Media, Inc.. ISBN-10: 0-596-52926-0, 2007.

Simonis, I. (ed.), OGC® Sensor Web Enablement Architecture, OGC Best Practices Document 06-021r4 Version: 0.4.0, http://www.opengeospatial.org/standards/bp, 2008.

Usländer, T. (ed.), Reference Model for the ORCHESTRA Architecture Version 2, OGC Best Practices Document 07-097, http://www.opengeospatial.org/standards/bp, 2007.

Usländer, T., The Growing Importance of Open Service Platforms for the Design of Environmental Information Systems, in: Proceedings of iEMSs, ISBN: 978-84-7653-074-0, Vol. 3, pp. 1628-1635, 2008a.

Usländer, T., Integration of Resource-Oriented Architecture Concepts into the OGC Reference Model. OGC Document 07-156r1, 2008b.

Usländer, T. (ed.), Specification of the Sensor Service Architecture, Deliverable D2.3.x of the European Integrated Project SANY, FP6-IST-033564, http://www.sany-ip.eu, 2009.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Web Information extraction for e-environment

Jan Dědek[a,b] **and Peter Vojtáš**[a,b]

[a] *Department of software engineering, Charles University, Prague, Czech Republic*
(jan.dedek@mff.cuni.cz, peter.vojtas@mff.cuni.cz)

[b] *Institute of Computer Science, Czech Academy of Sciences, Prague, Czech Republic*

Abstract: We will discuss possibility of using web information extraction methods for improving understanding eEnvironment relevant information on the web. Main contribution is in automated information extraction from web resources and annotation by an ontology.

Keywords: Web; Information extraction; Natural Language Processing.

1 INTRODUCTION AND MOTIVATION

Today a big amount of information is published electronically on the web. Good examples are pages of public institutions, which are publishing plenty of information on their web pages. This does not except information about the environment. In the Czech Republic for example the public right to information about the state of environment is guaranteed by the *Charter of Fundamental Rights and Basic Freedoms* (article 35/2) and the information is provided for example by *Ministry of the Environment* and by *local governments*.

This information is published in the form of natural language texts, which are suitable for human readers but not for computer processing. Computer processing of the information can be beneficial in many directions – statistical purposes, easy information search, integration of information from different sources, artificial intelligence automatic reasoning for new derived knowledge, automatic detection of contrary claims, complex visual presentation and publication, etc. All the benefits of machine understandable information sketched Tim Berners-Lee et al. [2001] in the famous article about the *Semantic Web*. Since then these ideas are being intensively developed within many activities of the *Semantic Web* foundation[1].

In the present paper we try to describe our method how the machine understandable data can be obtained form the web. We concentrated ourselves on the data and information relevant to the environment and we also present a practical experiment with extraction of information related to risks of environment's damage.

As described on the Figure 1 the web contains just partial information about environment and environmental risks. And just a part of it can be extracted by web information extraction tools but even though such information can make up interesting and important evidence. The fact that this evidence is kept in machine understandable form brings us all the advantages mentioned above.

The paper is structured as follows. Next section describes our extraction system in detail. Then section 3 is concentrated on the extraction method itself. The section 4 presents our experiment with environmentally relevant data and section 5 concludes the paper.

[1] http://www.w3.org/2001/sw/

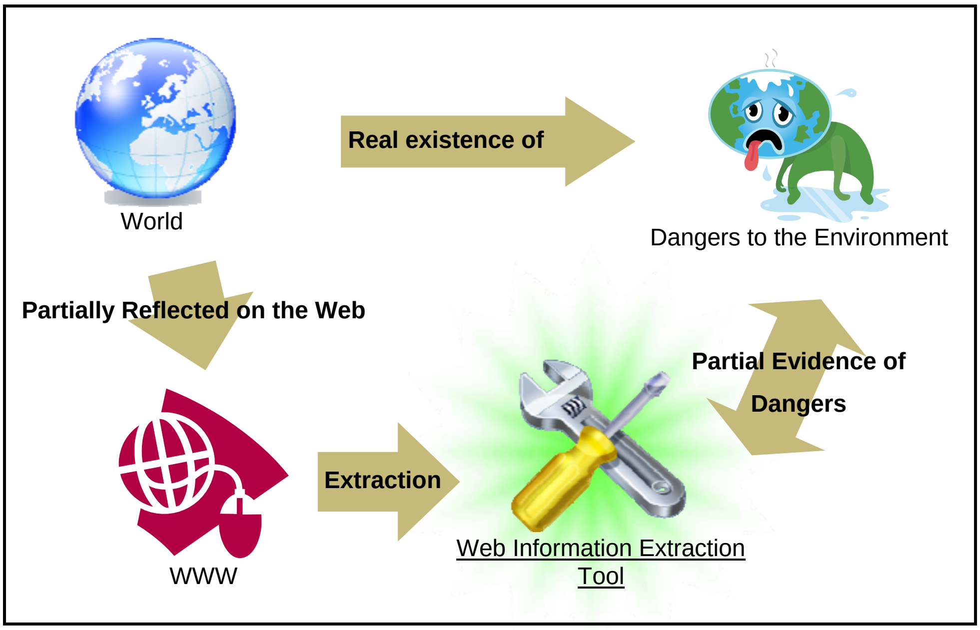


Figure 1. The Environment and Information on the Web

2 DESCRIPTION OF OUR TOOL / SYSTEM

We are developing a system, which produce machine (computer) processable information from the web. Our system captures text of web pages, annotates it linguistically, extracts data and stores the data in ontology data structure[2].

Our system covers an extraction process that starts on the web and ends in an ontology. This process consists of four steps. The Figure 2 describes them.

1) Extraction of relevant text
In this phase we have to extract the text from a page on the web. We use RSS feeds of the target web site. From the RSS we obtain URLs of particular articles (web pages) and we download them. From downloaded web pages we extract the desired text by means of a regular expression. This text is an input for the second phase.

2) Linguistic annotation
In this phase the linguistic annotation tools process the extracted text and produce corresponding set of linguistic trees representing the deep syntactic structure of individual sentences. We have used third party linguistic tools, which will be described in next section.

3) Information extraction
Our extraction method will be described in next section. This method uses the structure of linguistic trees and special extraction rules. The extraction rules define the target of extraction.

4) Semantic interpretation
This phase consists of transformation or conversion of the structured extracted data to the desired

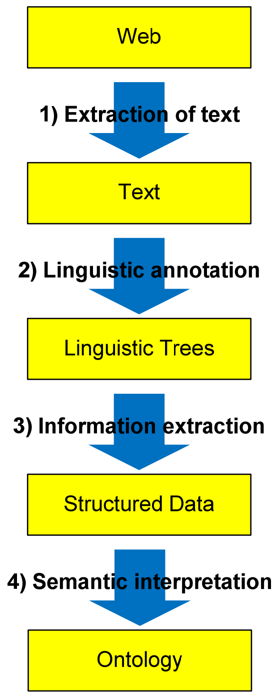


Figure 2. Data flow.

[2] Web Ontology Language http://www.w3.org/TR/owl-guide/

ontology format. It is quite important to choose suitable ontology that will properly represent semantics of the data. The interpretation expresses how to transform matching nodes of an extraction rule (and the available linguistic information connected) to the format of output ontology. Complexity of the transformation varies form simple (e.g. setting value of a data-type property to the value of some linguistic attribute) to complex. More details can be found in our previous work [Dědek and Vojtáš, 2008].

3 EXTRACTION METHOD

Our extraction system exploits set of linguistic tools published in [Hajič et al. 2006] and [Klimeš 2006]. These tools make tokenization, segmentation, morphological analysis, and linguistic parsing on so called *analytical* and *tectogrammatical* level of Czech. Although this linguistic analysis is addressed to the Czech language our extraction method is not limited to Czech. The project PEDT[3] uses the same linguistic theory for English and we plan to demonstrate usability of our method also with different linguistic approaches[4].

Our extraction method is based on extraction rules. These rules correspond to query requests of Netgraph[5] application. The Netgraph application is a linguistic tool used for searching through a syntactically annotated corpus of a natural language. Jiří Mírovský [2008] finished the development of Netgraph application (as his doctoral thesis) recently. Netgraph queries are written in a special query language. An example of Netgraph query will be described in next section.

The extraction works as follows: the extraction rule is in the first step evaluated by searching through a set of linguistic trees. Matching trees are returned and the desired information is taken from particular tree nodes.

Figure 3 shows linguistic (tectogrammatical) tree of sentence: "*Due to the clash the throat of fuel tank tore off and 800 litres of oil (diesel) has run out to a stream.*" (In Czech orignal: *Nárazem se utrhl hrdlo palivové nádrže a do potoka postupně vyteklo na 800 litrů nafty.*) Some of the nodes are emphasised – this should demonstrate how the extraction rule matches the tree.

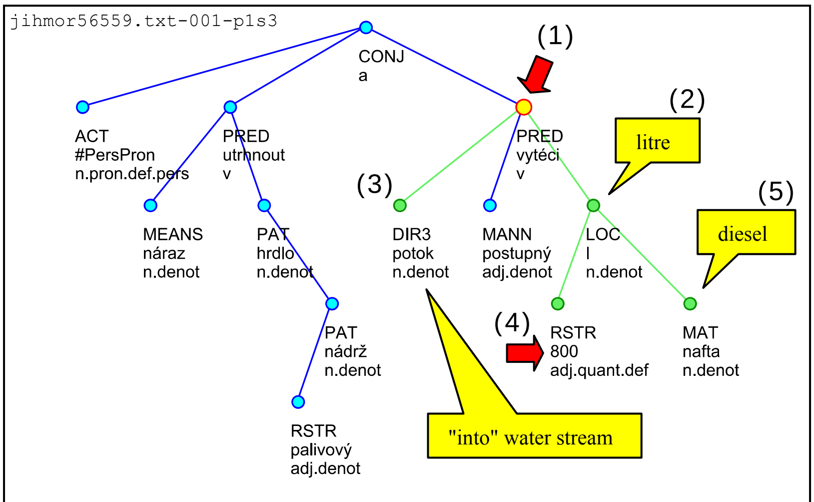


Figure 3. Example of linguistic tree.

[3] Prague English Dependency Treebank, http://ufal.mff.cuni.cz/pedt/

[4] *The Penn Treebank Project*, http://www.cis.upenn.edu/~treebank/ represents one of the suitable approaches. Its linguistic annotations are supported by many available automatic tools e.g. the *Stanford POS Tagger* and *Parser*, http://nlp.stanford.edu/software/

[5] http://quest.ms.mff.cuni.cz/netgraph/

4 EXPERIMENTS WITH ENVIRONMENTALLY RELEVANT DATA

For the purpose of the *TOWARDS eENVIRONMENT* conference we have made some experiments with acquisition of environmentally relevant information form the web.

We have used web reports from fire departments of several regions of the Czech Republic. These reports are written in Czech language and can be accessed through the web of *General Directorate of the Fire and Rescue Service of the Czech Republic*[6]. These reports describe interventions of fireman's units during different accidents (car accidents, fire accidents, etc.). These reports are rich in information, e.g. where and when a traffic accident occurred, which units helped, how much time it took them to show up on the place of accident, how many people were injured, killed etc. There is also information about endangered environment in the reports. Here we present an experiment with acquisition of environmentally relevant information from the fireman's web reports.

In the experiment we were interested in monitoring dangerous liquids (or materials) that have run out (spilled) during an accident. For this purpose we have designed an extraction rule that is depicted on the Figure 4. This rule consists of five nodes. Each node of the rule will match with some node in each suitable tree. So we can investigate the relevant information by reading values of linguistic tags of matching nodes. The fact, that this extraction rule match with some linguistic tree in most cases means that the tree (and corresponding sentence) deals with an amount of something which have run out (spilled). This is partially ensured by the node (1) – root of the extraction rule, which is restricted to contain one of the verbs "`uniknout`", "`unikat`" and "`vytéci`", which all have similar meaning to the English verbs *run out* or *spill*. From the other nodes we can find out the amount (node number 4), metric units (2) and material (5), which have run out (spilled) during an accident. And we can also identify the location where the dangerous material ended (expressed by the optional node number 3).

The extraction rule on the Figure 4 is designed manually by human expert, who has to be familiar with Netgraph application and linguistic formalism used. In [Dědek, Eckhardt, Vojtáš, 2008] we have presented an approach based on *inductive logic programming*. This approach makes it possible for almost unskilled user to tag relevant words in a sentence. Extraction rules are then learned automatically by the system.

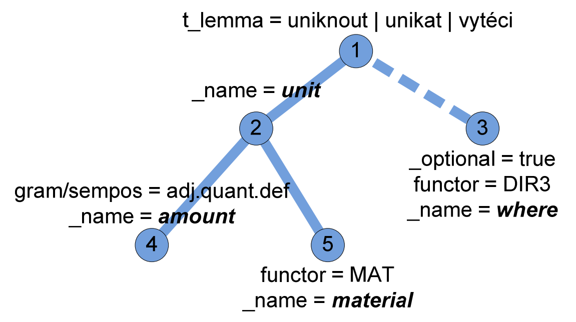


Figure 4. Example of extraction rule.

4.1 Results

Some example pieces chosen form the XML results are shown in the Figure 5. This example results contains four pieces of information extracted from four articles using the extraction rule on the Figure 4. Each piece of extracted information corresponds to a match of the extraction rule. Each match is closed in the `Match` element and each contains values

---

[6] http://www.hzscr.cz/

of some linguistic attributes closed inside the `Value` elements. Each value comes from some of the nodes of the extraction rule. Name of corresponding node is saved in the `variable_name` attribute of the `Value` node. The original text of matching sentence is closed inside the `Sentence` tag. We can see that the first matching sentence corresponds with the tree presented on the Figure 3 (with emphasized matching nodes).

Different materials (diesel, gear oil and some unspecified material – probably mentioned in related sentences) are showed up and different endangered types of location (water stream and soil) are identified in presented examples.

```
<QueryMatches>
  <Match root_id="jihmor56559.txt-001-p1s3" match_string="15:0,16:4,22:1,23:2,27:3">
    <Sentence>Nárazem se utrhl hrdlo palivové nádrže a do potoka postupně vyteklo na
     800 litrů nafty.</Sentence>
    <Data>
      <Value variable_name="amount" attribute_name="t_lemma">800</Value>
      <Value variable_name="unit" attribute_name="t_lemma">l</Value>
      <Value variable_name="material" attribute_name="t_lemma">nafta</Value>
      <Value variable_name="where" attribute_name="t_lemma">potok</Value>
    </Data>
  </Match>
  <Match root_id="jihmor68220.txt-001-p1s3" match_string="3:0,12:4,21:1,22:2,27:3">
    <Sentence>Z palivové nádrže vozidla uniklo do půdy v příkopu vedle silnice
     zhruba 350 litrů nafty, a proto byli o události informováni také pracovníci
     odboru životního prostředí Městského úřadu ve Vyškově a České inspekce
     životního prostředí.</Sentence>
    <Data>
      <Value variable_name="amount" attribute_name="t_lemma">350</Value>
      <Value variable_name="unit" attribute_name="t_lemma">l</Value>
      <Value variable_name="material" attribute_name="t_lemma">nafta</Value>
      <Value variable_name="where" attribute_name="t_lemma">půda</Value>
    </Data>
  </Match>
  <Match root_id="kralovehrad54765.txt-001-p6s5" match_string="1:0,7:1,8:2,13:3">
    <Sentence>Z kamionu uniklo zhruba 20 litrů látky.</Sentence>
    <Data>
      <Value variable_name="amount" attribute_name="t_lemma">20</Value>
      <Value variable_name="unit" attribute_name="t_lemma">l</Value>
      <Value variable_name="material" attribute_name="t_lemma">látka</Value>
    </Data>
  </Match>
  <Match root_id="moravslez50487.txt-001-p4s1" match_string="43:0,49:1,50:2,55:3">
    <Sentence>Hasiči po likvidaci požáru trávy asi na 25 metrech čtverečních ještě
     uklízeli společně s pracovníky Správy silnic Moravskoslezského kraje zhruba 15
     metrů silnice, na kterou vyteklo asi 40 litrů hydraulického oleje.</Sentence>
    <Data>
      <Value variable_name="amount" attribute_name="t_lemma">40</Value>
      <Value variable_name="unit" attribute_name="t_lemma">l</Value>
      <Value variable_name="material" attribute_name="t_lemma">olej</Value>
    </Data>
  </Match>
</QueryMatches>
```

litre

diesel

water stream

soil

other material

gear oil

Figure 5. Example of extraction result.

5 CONCLUSION

We have presented our system for extraction of information from Czech text on Web pages. Our system relies on linguistic annotating tools from ÚFAL[7] and the tree querying tool Netgraph. Our method is not limited to Czech and can be used with other languages and similar linguistic approaches. Although our system is still in development it can produce interesting results today.

We have shown an example how the environmentally sensitive information can be extracted from fireman's web reports and we are convinced that our method is applicable in any similar setting. As mentioned in our motivation: just partial information about the environment can be extracted from the web but even though such information can make up interesting and important evidence. The fact that this evidence is kept in machine understandable form brings us all the advantages of the semantic web technologies.

[7] Institute of Formal and Applied Linguistics in Prague, http://ufal.mff.cuni.cz/

ACKNOWLEDGEMENTS

This work was partially supported by Czech projects IS-1ET100300517, GACR-201/09/H057 and MSM-0021620838.

REFERENCES

Berners-Lee, T., Hendler, J. and Lassila, O., The Semantic Web. *Scientific American*, May 2001, 34–43.

Dědek, J., Eckhardt, A. and Vojtáš, P. Experiments with Czech linguistic data and ILP. *ILP 2008 - Inductive Logic Programming (Late Breaking Papers)*, Železný, F. and Lavrač, N. Prague, Czech Republic: Action M, 2008, 20–25.

Dědek, J. and Vojtáš, P., Computing aggregations from linguistic web resources: a case study in Czech Republic sector/traffic accidents. *Second International Conference on Advanced Engineering Computing and Applications in Sciences*, C. Dini, Ed. IEEE Computer Society, 2008, 7–12. http://www2.computer.org/portal/web/csdl/doi/10.1109/ADVCOMP.2008.17

Hajič, J. et al. Prague dependency treebank 2.0 cd-rom. *Linguistic Data Consortium LDC2006T01*, Philadelphia 2006.

Klimeš, V. Transformation-based tectogrammatical analysis of czech. *The 9th International Conference TSD 2006*, Lecture Notes In Computer Science 4188, 135–142. Springer-Verlag Berlin Heidelberg, 2006.

Mírovský, J. Netgraph – a Tool for Searching in the Prague Dependency Treebank 2.0, Doctoral Thesis (2008), *Institute of Formal and Applied Linguistics*, Faculty of Mathematis and Physics, Charles University in Prague, 2008. http://quest.ms.mff.cuni.cz/netgraph/publications.html

Session 2

Building Shared Environmental Information System, Europe's backbone for eGovernment environmental services

Organized by **Jiří Hradec**, Timo Mäkelä, Meropi Paneli, Thomas Pick, Giorgio Saio and Chris Steenmans

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunites of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pilmmann, I.Holoubek, T. Bandholz (eds.)
Masaryk University, 2009

Data harmonization of environmental variables: from simple to general solutions

Olivier P. Baume[a], Jon O. Skøien[b], Florence Carré[c], Gerard B.M. Heuvelink[a] and Edzer J. Pebesma[d]

[a] *Department of Environmental Sciences, Wageningen University, the Netherlands (olivier.baume@wur.nl; gerard.heuvelink@wur.nl)*

[b] *Department of Physical Geography, University of Utrecht, the Netherlands (j.skoien@geo.uu.nl)*

[c] *Land Management & Natural Hazards Unit, DG JRC, Ispra, Italy (florence.carre@jrc.it)*

[d] *Institute for Geoinformatics, University of Münster, Germany (edzer.pebesma@uni-muenster.de)*

Abstract: European environmental databases can be merged if regional networks or experimental campaigns target the same variable. In practice, measurements contain biases especially when compared from one network to another, or from one experiment campaign to another. Therefore merging is sensible only when data are harmonized. Harmonization of European environmental databases is often meaningful for decision making, and sometimes it is crucial. In this paper, a solution for data harmonization, developed under the INTAMAP FP6 project, is presented. The INTAMAP FP6 project is currently developing an interoperable framework for real time automatic mapping of critical environmental variables by extending spatial statistical methods. The harmonization procedure takes only additive biases into account. The bias is estimated and removed before final interpolation. We describe two applications of harmonization at a European level. The original motivation for INTAMAP came from the field of radiology, where a European real-time database has been active for all member states over a decade. This so-called EURDEP database contains heterogeneities because national networks measure and report data following different devices and processes. As an example, monthly averaged gamma dose measurements across eight European countries are harmonized. The second application deals with the soil Carbon/Nitrogen (CN) ratio in Europe. The CN ratio of the forest soil cover is one of the best predictors for evaluating soil functions, such as required in climate change issues. Although samples were analyzed according to a common laboratory method, preliminary statistical analyses showed that laboratories introduced errors in the measurements and should be taken into account. Results from both applications are discussed to continue the work of harmonization.

Keywords: Measurement error; Sensor calibration; Geostatistics, Uncertainty, Environmental model.

1. BACKGROUND

A problem with international environmental databases is that values reported on a specific target variable may not have followed the same data process from one network to another, or from one campaign to another. Examples include the use of different measuring devices, the use of different standards, or the analysis of samples in different laboratories. Such discrepancies in a database may lead to wrong decisions in policy making: mapping environmental variables across borders may lead to wrong detections of threshold values at which remediation actions are required; mapping environmental variables at a European scale may lead to wrong analysis and model building of the target variable.

In environmental assessment, while standardization implies the definition of a common way to measure a specific target variable, harmonization is a bottom-up approach for gathering different standards. Standardization is generally accepted as an obvious solution (see for instance Wagner et al. [2001], Schröder at al. [2006]), although standardization can become very costly when national databases already exist for many years. In the extreme case of several historical campaigns, standardization is not an option for analyzing legacy data. As a more flexible alternative to standardization, several authors from the field of geostatistics suggest harmonization procedures to handle heterogeneous databases.

Harmonization modelling makes a clear distinction between the target variable that one wants to map, and the measured variable that is in fact reported from a particular network or measurement campaign. Fassó et al. [2007] applied a space-time kriging model for air quality mapping in the region of Lombardy, which includes additive and multiplicative biases between two measurement networks of PM10. Brenning et al. [2008] analysed the discontinuity errors between field measurements to map soil conductivity from field scale to landscape scale. Skøien et al. [2009] used a line kriging method to identify discontinuities at country borders among 30 networks of the European radioactivity exchange database. Baume et al. [2007] applied a universal kriging model to filter additive biases between national networks. The latter was extended to a Bayesian setting in order to include prior knowledge on the biases between networks (Baume et al. [2008]).

All harmonization methods have been applied to specific examples. In this paper we consider the method developed by Baume et al. [2008], which may be applied with or without prior information on the biases. We apply the method to two different examples. First an application on radioactivity exposure aims at removing biases between national networks in order to process automatic mapping of harmonized data. The example further uses the network gaps estimated by Skøien et al. [2009] as prior information in the model. The second example targets soil quality data of the soil Carbon/Nitrogen (CN) ratio from Europe. Harmonization of several national campaigns may lead to a better understanding of the influence of tree type on CN ratio values.

2. METHODS

The main principle of harmonization is to distinguish a common target variable Z defined for the whole database, from the measured variables Y_i, which are indexed for each specific network or measurement campaign. A way to denote the relationship is to use some general calibration functions F_i so that

$$Z(\mathbf{S}_i) = F_i^{-1}(Y_i(\mathbf{S}_i)), \tag{1}$$

where the $\{\mathbf{S}_i\}$ are the measurement locations of the specific network or measurement campaign i. The calibration functions F_i are device-dependent. For simplicity we assume that one additive bias b_i per network or campaign i provides sufficient harmonization. Thus, equation (1) reduces to

$$Z(\mathbf{S}_i) = Y_i(\mathbf{S}_i) - b_i. \tag{2}$$

2.1 Geostatistics and harmonization

There are two main methods to estimate the biases b_i. The first method is to indirectly consider the local differences between measurements from neighbouring networks or campaigns, and to transform the estimated differences into biases (see Brenning et al. [2007], Skøien et al. [2009]). The second method is to directly consider the presence of biases in an overall linear model by means of a universal kriging model (Baume et al. [2007]). In this paper we consider the second method.

Universal kriging is a geostatistical method for mapping environmental variables that may include some explanatory variables $\{X_k\}$ as well. For example, the background radioactivity level (gamma dose for global radioactivity exposure) is influenced by the altitude of the measurement and by soil type. In the general case, we relate the target variable Z with K explanatory variables X_k through K coefficients A_k to be estimated. Writing the relationships in terms of random variables, and introducing residuals R_Y and R_Z, the overall linear model of universal kriging for harmonization writes:

$$\begin{cases} Y_i(\mathbf{S}_i) = & Z(\mathbf{S}_i) + b_i + R_{Y_i}(\mathbf{S}_i) \\ Z(\mathbf{S}_i) = & \sum_{k=1}^{K} X_k(\mathbf{S}_i) \cdot A_k + R_Z(\mathbf{S}_i) \end{cases} \tag{3}$$

.

We assume that the stochastic residuals R_Y and R_Z have zero-mean and that residuals of the true state variable R_Z have a constant covariance structure in the whole domain of study.

2.2 A harmonized database for mapping and modelling

We solve the system (3) by means of the least squared difference method (see Baume et al. [2007] for a detailed description). In a first step, we obtain an estimate of the biases $\{b_i\}$, an estimate of the linear relationships A_k between the target variable and its explanatory covariates, and an estimate of the covariance of the true process R_Z. We use a simple iterative process from ordinary least squares estimation to general least squares, until convergence. In a second step, as a result of universal kriging, it is possible to directly compute an optimized interpolation where the biases are removed.

We call this interpolation procedure harmonized kriging. Optionally some information can be added on the model prior to the estimation procedure. The inclusion of prior information on the biases has been reported by Baume et al. [2008] in the context of radioactivity exposure. In harmonized kriging, the inclusion of prior information is guided by the fact that in linear model (3), multicollinearity among covariates and biases is a source of estimation impairment and misinterpretation.

In a simple framework that only includes additive biases b_i in model (3), one problem that remains with harmonized kriging is the choice of a reference. By default, the biases are computed such that they sum up to 0, but alternatively, if the measurement standard of one of the networks or campaigns can be elected as a reference, its bias may be set to 0.

3. APPLICATIONS

Harmonization is anticipated to give more accurate decisions because the harmonization model (3) makes an explicit distinction between the measured variables and the true state variable. Kriging interpolation maps provide a probability distribution of the target variable at each interpolated location that eventually lead to the probabilistic computation of threshold detections. Without a harmonization procedure in the analysis of a database crossing borders, the risk of false decision is greater.

3.1 Radioactivity exposure and risk assessment

We take the example of bias estimation in gamma dose rate data (physical unit nSv/h) from the European Data Exchange Platform (EURDEP – http://eurdep.jrc.it/). This database is used for radioactivity background level estimation and for early warning situations. Monthly averages of December, 2006 were extracted for eight countries – Austria (AT), Belgium (BE), Switzerland (CH), Czech Republic (CZ), Germany (DE), Italy (IT), Luxemburg (LU) and Netherlands (NL). Within each of these countries the data are assumed to belong to one network only. Different networks may use different probe types and data processing before the upload to the EURDEP database. For instance, Germany subtracts the self-effect of the probes whereas not all other countries do. Despite an attempt to collect such information from each country (Bossew et al. [2007]), there are still differences between national networks that are not accounted for. As a consequence, the harmonization can only be based on the national membership of the measurements.

Following model (3), we assumed one bias per country and included two covariates in the linear model of the target variable *Z*. The first covariate is elevation because elevation influences the amount of terrestrial radiation detected: secondary cosmic radiation raises exponentially with altitude hence augmenting background gamma dose rate level (Wissmann et al. [2008]). As most measurement sites are located at low altitudes, non-symmetric distributions appear for mountainous countries. The second covariate is soil type. For instance, radioactivity in volcanic regions is generally higher. Lowest background radioactivity can be found mostly in alluvial plains.

To improve harmonization, the harmonization model (3) was completed with prior information on the biases (see Baume et al. [2008]). These prior values correspond to local estimation of biases along borders. Local estimates of the biases along borders - and respective standard deviations – were obtained with the method developed by Skøien et al. [2009]. As these local estimates are less sensitive to covariates, they can be considered sensible prior information. The prior values of biases are given in Table 1.

3.2 Carbon/Nitrogen ratio for evaluating soil functions

As an indicator of soil mineralization processes, the CN ratio of forest soils is one of the best predictors for evaluating soil functions such as biomass production and carbon storage capacity of forest soils. When integrated to risk assessment, these functions can serve for modelling scenarios of soil sustainability with climate change issues (e.g. gas fluxes emissions, biofuel production). For instance, for a soil having a relative high CN ratio, the mineralization process tends to be slower and the weak leaching of nitrogen results in a weak quantity of N gas fluxes emission.

The CN ratio is strongly dependent on the forest species and management, and on environmental factors (Burke et al. [1989]) such as climate, relief, soil type and parent material. Furthermore, since the forest management is done locally, the CN ratio variability has to be analysed locally. After a preliminary data analysis (Carré et al. [2008]), data harmonization appears an important step to level out main gaps between local sample analysis campaigns. Analysis of soil cover samples were analysed by several laboratories, each having somewhat different procedures and circumstances. Therefore laboratory origin is considered as the origin of bias in the dataset. Most countries centralised the sample analysis in one laboratory except for Germany, where analysis was centralised per region. In the CN ratio application, the model of the target variable *Z* included two covariates. First, each measurement sample was related to the percentage of two complementary tree cover types: coniferous cover and broadleaves cover. Second, the model included the pH as a continuous covariate. No prior information on the biases was available in this application.

4. RESULTS AND DISCUSSION

4.1 Influence of prior information on bias estimation in Radioactivity exposure

A comparison of bias estimates in the radioactivity case given in Table 1 shows that both the local method ("Prior values", Skøien et al. [2009]) and harmonized kriging ("Without prior") yield similar results. Table 2 gives the "Posterior with prior" estimates of the country biases. Standard deviation values associated to the prior estimates (Table 1, column 4) correspond to the inverse of the weight given to the prior values to estimate the posterior values. In this example, prior values have larger standard deviations than harmonized kriging estimates because local estimation of biases rely only on the measurements close to the borders. In the extreme case of Italy where the network is sparse, the standard deviation is the largest.

Table 1. Radioactivity exposure – gamma dose: network bias estimates and associated standard deviations for harmonization with and without prior information (*nSv/h*).

Country	without prior	*Standard deviation*	Prior values	*Standard deviation*	Posterior with prior	*Standard deviation*
Austria (AT)	-18.31	*2.04*	-15.00	*3.24*	-18.29	*1.96*
Belgium (BE)	10.76	*2.66*	14.15	*2.95*	12.25	*1.93*
Switzerland (CH)	10.33	*2.98*	11.23	*2.55*	10.36	*1.93*
Czech Republic (CZ)	-0.63	*2.87*	2.18	*3.95*	0.10	*2.27*
Germany (DE)	-5.88	*1.43*	-3.52	*2.60*	-5.20	*1.28*
Italy (IT)	-2.81	*4.41*	-18.97	*12.14*	-6.78	*3.76*
Luxembourg (LU)	16.02	*4.08*	17.96	*5.49*	16.18	*3.15*
Netherlands (NL)	-9.49	*2.09*	-8.05	*2.85*	-8.62	*1.64*

4.2 Interpretation of harmonized kriging in the Carbon / Nitrogen ratio mapping

As no prior information was included in the CN ratio case study, we plotted the histogram of the correlation level between natural covariates (cover type and pH) and the membership of the measurements (related to countries and regions) to evaluate the risk of impairment in the estimation process (Figure 1). Correlation values are low with absolute values lower than 0.4. The majority of correlation absolute values in the model are lower than 0.1. Table 2 presents the bias estimates and their standard deviation. Figure 2 compares the interpolation map without correction of the biases (mentioned as "ordinary kriging") and the interpolation with biases removed (using "harmonized kriging").

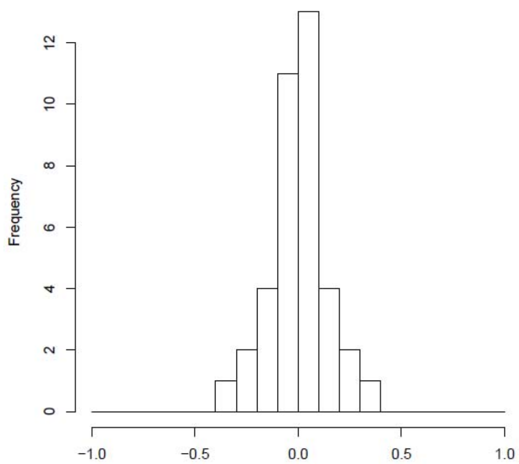


Figure 1. Carbon/Nitrogen ratio: histogram of correlation coefficients between natural covariates and country / region membership.

CN ratios are overestimated in Belgium, Finland and Latvia, whereas Great-Britain and Poland for example were underestimating the CN values (Table 2). Lowest standard deviations are found in largest regional datasets. Harmonization decreases differences between countries. However, the harmonized map still shows hotspot values in France, Latvia, Germany and Poland (see Figure 2(b)). Hotspots are due to the presence of coniferous forests and low pH (acid soils like podzols), for which the decomposition of organic matter is slow (the CN ratio is then high). pH and forest species allowed to take the soil environment into account for correcting the CN ratios.

Table 2. Carbon / Nitrogen ratio: country/laboratory biases. Data were analysed by one laboratory in each country, except for Germany.

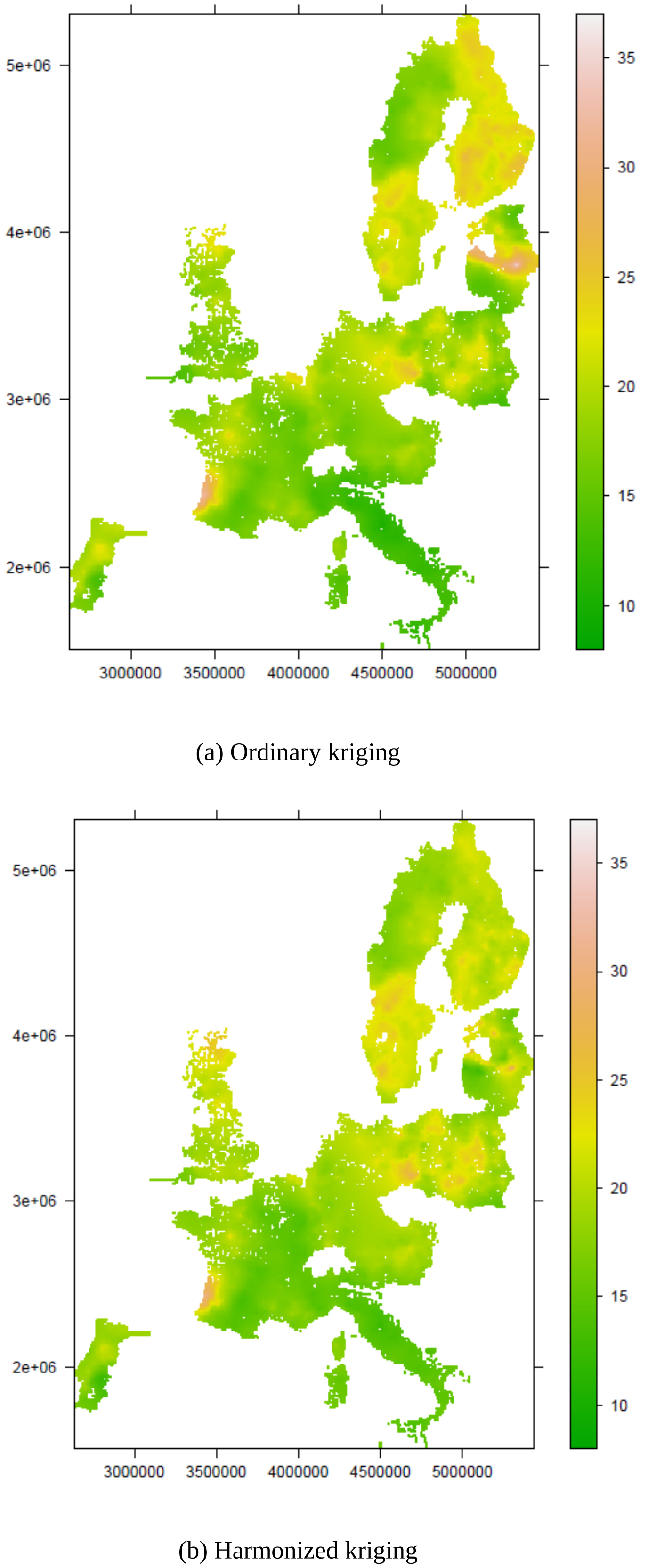
Country (lab. code)	Bias estimate	*Standard deviation*
Austria	-1.2	*0.8*
Belgium	2.6	*2.0*
Estonia	-1.5	*0.9*
Finland	2.4	*0.5*
France	0.9	*0.5*
Germany (401-Lab01)	-2.1	*1.2*
Germany (404-FHE)	0.7	*1.2*
Germany (408-lab_0)	-1.2	*1.6*
Germany (410-lab_g)	0.5	*1.2*
Germany (411-1)	-0.7	*1.6*
Germany (412-Lab_4)	1.5	*2.3*
Italy	-2.0	*0.6*
Latvia	7.5	*1*
Lithuania	-2.8	*1.1*
Poland	-1.7	*0.5*
Portugal	1.3	*0.8*
Slovenia	-0.7	*1.5*
Sweden	-1.0	*0.5*
Great-Britain	-2.4	*0.6*

The harmonized map also shows spatial gradients in the values, notably in England (from North to South), France (from West to East) and Scandinavia (from North to South). These gradients can mainly be explained by bioclimatic gradients which have an impact on the decomposition rate of the organic matter (humidity and high temperature increase the rate of decomposition or decrease the CN ratio). The harmonized kriging map has a stronger correlation than the ordinary kriging map with variables that have not been taken into account in the harmonization process.

5. CONCLUSION

We proposed a geostatistical modelling solution to the harmonization problem of environmental databases with additive biases. We applied the harmonization methodology to two different datasets. The first example was taken from the problem of radioactivity exposure assessment and shows that the estimation of additive biases in a linear model may yield results similar to a more local procedure, such as the method developed by Skøien et al. [2009]. The second example, dealing with the Carbon/Nitrogen ratio of forest soils indicates that the harmonized map is more correlated to important related factors such as bioclimatic synthetic variables.

As a conclusion, the simple solution that we propose seems to work in different contexts, especially when the number of networks or laboratories is large. The model can be improved with prior information, which gives a guarantee to avoid estimation impairment due to multicollinearity in the linear model. The method will be validated for the Carbon/Nitrogen ratio case study in July 2009 after a second analysis of the measured data by one central laboratory (the French laboratory that had already analyzed the French data).



(a) Ordinary kriging

(b) Harmonized kriging

Figure 1. Carbon/Nitrogen ratio maps of forest soils for parts of Europe.

ACKNOWLEDGMENTS

This work is funded by the European Commission, under the Sixth Framework Program, by the Contract N. 033811 with the DG INFSO, action Line IST-2005-2.5.12 ICT for Environmental Risk Management. The views expressed herein are those of the authors and are not necessarily those of the European Commission.

REFERENCES

Baume, O., Skøien, J.O., Heuvelink, G.B.M. and Pebesma, E.J., Geostatistical approach to data harmonization. Presented at: *statGIS2007*, Klagenfurt, Austria, 2007.

Baume, O., Skøien, J.O., Heuvelink, G.B.M., and Pebesma, E.J., Data harmonization with geostatistical tools: a Bayesian extension. Presented at: *Geostats2008*, Santiago, Chile, 2008.

Bossew, P., De Cort, M., Dubois, G., Stöhlker, U., Tollefsen, T. and Wätjen, U., AIRDOS, Evaluation of existing standards of measurement of ambient dose rate; and of sampling, sample preparation and measurement for estimating radioactivity levels in air. TREN/NUCL/S12.378241, JRC ref. 21894-2004-04 A1CO ISP BE, 2007.

Brenning, A., Koszinski, S., and Sommer, M., Geostatistical homogenization of soil conductivity across field boundaries, *Geoderma*, 143 (3-4): 254-260, 2008.

Burke, I.C., Yonker, C.M., Parton, W.J., Cole, C.V., Flach, K.D., and Schimel, S., Texture, climate, and cultivation effects on soil organic matter content in U.S. grassland soils. *Soil Sci. Soc. Am. J.*, 53:800-805, 1989.

Carré, F., Jeannée, N., Casalegno, S., Lemarchand, O., Reuter, H.I., and Montanarella, L., Mapping the CN ratio of the forest litters in Europe – lessons for global DSM**.** Presented at: *Logan2008*, Salt Lake City, United States, 2008.

Fassó, A., Cameletti, M., and Nicolis, O., Air quality monitoring using heterogeneous networks, Environmetrics 18: 245-264, 2007.

Schröder, W., Pesch, R., and Schmidt, G., Identifying and closing gaps in environmental monitoring by means of metadata, ecological regionalization and geostatistics using the UNESCO biosphere reserve Rhoen (Germany) as an example, *Environmental Monitoring and Assessment*, 114: 461-488, 2006.

Skøien, J.O., Baume, O., Pebesma, E.J., and Heuvelink, G.B.M., Identifying and removing heterogeneities between monitoring networks. *Environmetrics*, accepted, 2009.

Wagner, G., Desaules, A., Muntau, H., Theocharopoulos, S., and Quevauviller, P., Harmonization and quality assurance in pre-analytical steps of soil contamination studies - conclusions and recommendations of CEEM Soil project, *The Science of the Total Environment*, 264: 103-117, 2001.

Wissmann, F., Rupp, A., and Stöhlker, U., *Characterization of Dose Rate*, Zeitschrift Kerntechnik, in Press, 2008.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Abstract:

Waste Management using the GS1 Standards

Jim Bracken[a]
[a] *GS1 Ireland (jim.bracken@gs1ie.org)*

Since the 1970s, the EU has prioritised environmental protection as a key task for all member states and has set up a legal framework covering waste management and reporting. By adopting directives, which are to be transposed individually by all member states, the EU aims at creating innovation and new business opportunities. In 2006, a new directive (2006/12/EC) was approved, including "the principle of manufacturer's responsibility" and the "Extended Producer Responsibility" (EPR). Now manufacturers, importers or traders - depending on who is bringing products into the European market - are motivated to behave ecologically within their supply chains by no longer passing on the costs of waste and its disposal to governmental authorities.

Key issue is the lack of standards for the identification of:

- stakeholders
- waste produce
- waste classification

Also missing are:

- tracking and tracing systems
- efficient information sharing between various participants

Furthermore, waste management is becoming an increasingly global issue, neither national nor European in scope only.

Using GS1 Standards in combination with existing EU and national legislation can help in solving these drawbacks. For that purpose, waste supply chains could be organised with the following elements:

- all participants identified by a Global Location Number (GLN)
- services of those participants (such as: collecting, sorting, disposing, recovering, recycling, etc.) separated by using (serial) Global Trade Item Numbers ((S)GTIN)
- Attached to the GTIN: waste information using the nomenclature of the EU and national product identifier and the European Waste Catalogue (EWC)
- Information sharing by using EDI/ EANCOM® - Standards
- all item and flow data is stored in a data pool structured via Global Data Synchronisation Network (GDSN) data requirements
- combined supply chains for goods and waste

By applying trusted and proven technologies as the GS1 system the same standards as used for trade items can be deployed and utilised for Waste Management. These standards are neutral and user driven, updated permanently in more than 120 GS1 Member Organisations all over the world.

Croatian Environment Information System Conceptual Model

J. Butuči[a], H.Mesić[b]
[a]*Croatian Environment Agency, Trg maršala Tita 8, 10 000 Zagreb, Croatia*
(jasna.butuci@azo.hr)
[b]*Croatian Environment Agency, Trg maršala Tita 8, 10 000 Zagreb, Croatia*
(hana.mesic@azo.hr)

Abstract: Following SEIS principles, Croatian Environment Agency (CEA) has prepared By-law on Environment Information System. This legal instrument has been adopted by Government in June 2008 (Official Gazette 68/08) setting up structure, content, format, functionality and Environment Information System (EIS) maintenance manner, together with proscription of data flow mode (delivery terms, means and obligations). Responsibility for by-law implementation lay to the CEA.

In that term, as institution responsible for establishment, conducting, development, coordination and maintenance of EIS, in September 2008 CEA has compiled detailed Croatian Environment Information System Development Programme for period 2009-2012 (Programme).

Within this Programme *Conceptual Model for Croatian Environment Information System* was set in the way that assure guidelines for establishment of not existing data flow systems, but also for improving and customization of existing data flow and information systems in line to common EIS, by using modern IT tools such as Internet and satellite technology. The aim is to move from “manual/paper” or “reporting on call” manner to the data flow system organized in the way that assures constant data availability nearest to the data source, on an open and transparent way. EIS is anticipated as decentralized but integrated information system accessible via CEA Internet portal, organized in the common “information provide/use/share way” and based on data/information providers network.

Among general principles and guidelines, *Conceptual Model* presented in Programme defines competent institutions and obligated entities-data providers, divided by 4 thematic groups with 11 basic thematic areas and 42 subareas as per by-law on EIS. Further, information on existing monitoring and data flow systems, databases and information systems in place are recognized and listed.

As per general Government policy, *Conceptual Model* presented in Programme preparation was conciliate with the Central State Administrative Office for E-Croatia.

***Keywords**:* Environment information system; Environment data; Data flow; Data access; Data exchange.

1. INTRODUCTION

For years and in some cases decades, certain segments of the environment have been observed and data has been collected in the Republic of Croatia. The data relevant for the field of the environment have so far been conducted according to law regulations, under the jurisdiction of different bodies of the state administration – ministries, counties, local government and state institutes. Data bases were developed coherently with historical practice, for specific areas in scientific-research and other institutions, according with their fields of work and specific regulations. Legal regulations did not include to some extent the use of information systems nor regulated standards of information technology, unlike the countries of the European Union, which have even predefined these issues on the level of individual segments, which made their legal integration difficult.

The usefulness expectancy of the data acquired and the awareness of the need for a better information flow are much higher today. The tools for acquiring rich information were expensive so far, and the tools for analysis and spreading data didn't exist during the 1980's and even in the 1990's. But in the beginning of the 21st century, these tools and the connection possibilities of the digital age allow us to easily collect, distribute and act on the base of the information, which has a great role in the process of making strategic decisions as instead of shallow and folklore data, there is accurate data for observing the state of the environment.

By adopting the By-law on Environment Information System (By-law in further reading) Republic of Croatia has created a basic legal frame for creating a unique system. By-law gives clear pointers to the obligated legal entities-data providers, which adds to the decrease of the risk of further separation and development of separate incompatible systems. With the By-law the content, methodological bases of the system, duties, methods of delivering environment data for the needs of the system and the way the environment data is used have been prescribed. Furthermore, the need for the possibility of data collection and sharing the information and data processed and analysed according to the international and European methodology, meaning the possibility of trading data about the environment with other similar existing systems on the level of the European Union is also proscribed. By-law regulates the creation of the Croatian Environment Information System Development Programme (Programme in further reading) in which organisation, the method of coordinating and maintenance of Environment Information System (EIS in further reading), gives the list of obligated entities-data providers and other operational activity.

With this systematic approach, the Republic of Croatia has defined a path for creating a unique EIS which connects all existing data and information about the environment, acknowledging the principles:

- of editing information closest to the primary data source,
- acquiring the information once and distributing it among other participants of the system for various purposes,
- ensuring easier approach of the information to the end user, primarily the bodies of the government of the Republic of Croatia for the needs of the direction of the environment protection policy and monitoring of its implementation, as well as for scientific, expert and wider public needs.

In this work, the view of the Conceptual Model for Croatian Environment Information System is given from two aspects: the concept of the system itself and the concept of its conducting.

2. CONCEPTUAL MODEL

2.1. Conceptual Model of the System

Following the latest practice in European Union with indicator approach and SEIS principles, EIS is designed as a decentralised but integrated information system, available via an Internet portal, based on information and the data provider's network, defined by the National list of indicators (NLI in further reading).

NLI is produced taking into account regulations, reporting obligations, international contracts and respecting the specific national needs. In 2008 Croatian Environmental Agency (CEA) has completed lists of indicator for total of 15 thematic areas, meaning data sheets were made for altogether 266 indicators. At the same time, in the line with Environment protection law (Official Gazette 110/07), NLI is the base for reporting, especially for the State of the Environment Report.

EIS conceptual model consist of integrated information systems of all available data, data sets and environmental information via following individual systems, structured in four EIS basic groups:

1. Environment components
   - Air quality information system
   - Inland waters information system
   - Marine information system
   - Nature protection information system
   - Soil information system

2. Sectors pressures
   - Waste management information system
   - Agriculture and forestry information system
   - Industry and energy information system
   - Traffic and tourism information system

3. Health impact
   - Health and safety information system

4. Society responses
   - General environment topics information system.

Within these main basic groups content is defined through thematic areas and sub-areas. For each thematic area and/or sub-area, an information system is established as a part of the integral EIS. The basic groups with belonging thematic areas and sub-areas are shown in Table 1.

Table 1. Basic EIS groups, areas and sub-areas

BASIC EIS GROUPS	THEMATIC AREAS	THEMATICS SUB-AREAS
Environment components	Air	Climate change
		Ozone layer
		Air quality
	Inland waters	Inland waters quantity
		Inland waters quality
	Marine	Marine ecosystem
		Coastal environment
	Nature	Biodiversity
		Landscape diversity
		Protected nature
		Genetic modified organisms
	Pedology and lithosphere	Soil
		Use of mineral raw materials
		Land cover, use and change
		Geological valuable phenomena, objects and structures
Sectors pressures	Waste	Waste management
	Agriculture and forestry	Agriculture
		Forestry
		Fishery and aquaculture
	Industry and energy	Pollutant registry/Installations and industry
		Industrial and ecological accidents
		Chemicals
		Energy
	Traffic and tourism	Traffic and transport
		Tourism and recreation
Health impacts	Health and safety	Civil protection
		Noise
		Population and households
		Ionisation radiation and nuclear safety
		Light pollution
		Natural disasters (drought, earthquake, fire, flood, etc.)
		Pollution impact on health and life quality
		Socioeconomics data
Society responses	General environment topics	Environment protection documents
		Environment policy implementation
		Sustainable development and cleaner production

		Scientific and expert data
		Real, methodology and documental data
		Instruments and policy measures (economic instruments, inspections, environment impact assessment, etc.)
		Legal acts
		Environment protection participants
		Spatial plans

2.2. Indicator approach

National list of indicators defines basic data groups within 42 topics that are necessary for indicator calculation. Each of the basic data group needed for the calculation of indicator consists of multiple attributes. Individual attributes are needed when filling in requests for quality insurance and quality control, data flow recognition or other criteria such as recognising the responsible institution, frequency of measurements etc.

In the case of all attributes for individual data groups are the responsibility of a single institution, such data should be easily accessible, fulfilling prescribed measures of security and quality control of data. Availability and control of quality of the data with a greater number of attributes, which are the responsibility of different institutions becomes more difficult.

The nonexistence of certain attributes needed for particular data as well as the fact that by definition then there is no responsible institution that would be the source of these attributes is a special issue.

When establishing EIS it is crucial to recognise the attributes of individual data and defining the necessary data flow with the purpose of creating an efficient relational connection between data base sub systems of the EIS.

On the level of the group of attributes containing certain data required for the indicator calculation, different “membership” of attributes in individual data groups is recognised. This “membership” is recognized on the level of the data itself and three principles of data flow can be defined as follows:

- Data for calculating a single indicator is used for the needs of calculating another indicator in the same thematic area and/or sub area
- Data for the calculation of a single indicator is used for the needs of calculating another indicator in a different area and/or sub area (the principle of data flow is given in Figure 1.)
- Data for calculating one indicator used only for that purpose.

For establishing EIS it is necessary to recognise all data groups, their attributes, data flow, record format and quality control method in which all the participants of system establishment are to be included.

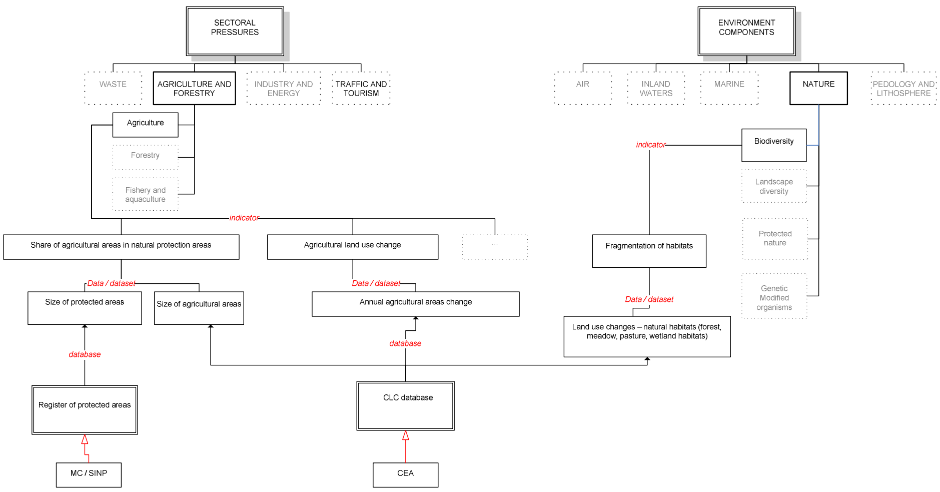


Figure 1: The principle of data flow used for the calculation of two indicators from different areas

2.3. Conceptual Model of the Conduction

Due to its complex nature, EIS development is planned to be build gradually on all administrative and other obligatory entities level. That comprehends considerable initially financial input for institutional strengthening crucial for technology changes (IT equipment, education, etc).

Organisation of EIS establishment and conducting activities is based on existing and newly adopted national legislation as well as on similar regional and European experiences, footing from cooperation between numerous dislocated, independently established, harmonised and connected thematic areas information systems.

According to proscribed obligations, Croatian Environment Agency is responsible for establishment, development, conduction, coordination and maintenance of the unique Environment information system.

Information system obligatory for legal entities have to ensure Operational plans drafting and conduction, nominate the representative(s) to NLI committee, nominate responsible persons for the information system establishment under their system/sub system jurisdiction, as well as to produce the protocols and measures for data quality control assurance.

Conduction concept is planned to be done through four phases:

PHASE I.	Setting up the technical background for system establishment
PHASE II.	Development and/or connection of information systems and databases of certain thematic areas and sub-areas in to EIS
PHASE III.	Technical and security maintenance and reporting
PHASE IV.	Implementation and education

For each phase the Programme defines specific goals and measures together with activities needed to reach those goals.

For construction of particular parts of the system and for their maintenance the performers are assigning. Performers are entities (institutes, faculties, companies) and experts (individuals).

3. PREDISPOSITIONS

Despite the strong political and social will to embrace the Conceptual Model of unique EIS, there are certain drawbacks in the process of its development. The value of success in this moment depends mostly on the possibility of insuring the organisation’s solutions and main elements of infrastructure, and of course last but not the least to ensure the sufficient financing.

Basic requests for Programme implementation require:

- harmonised work on project preparation,
- defining priorities for the projects implementation by using synergy effects in projects performance,
- defining individual information systems,
- defining of the bearer and the obligators,
- recognition of the primary data groups availability,
- possibility of long term planning for monitoring of certain environment component and/or pressure source as well as assurance of the precondition for predicted activities implementation,
- possibility of time and finance planning for individual information system projects,

- possibility to estimate the expenses needed for gathering the environment monitoring data.

Special demands for Programme implementation are:

- assurance of the structure applicable to assure growth and upgrade of the existing system for sub-areas not included in this four year Programme,
- assurance of complementary interconnections (taking over parts of projects/solutions from one project to another),
- optimal and rational use of overall resources (financial funds, experts etc.),
- assurance of project's implementation synchronisation.

4. CONCLUSION

The fact that Republic of Croatia did not participate, because of historical reasons, in production of vertical business solutions, which were already active for three decades in the computer industry, and legal obligations were not in place to define information systems, gives us certain advantage.

Conceptual Model implements the necessity of upgrading and integrating of already developed and implemented system, development of nonexistent systems, and by all means establishing of new law regulations where needed.

This *Conceptual Model* follows the needs to simplify the process of reporting and ensure the execution of requested obligations. Furthermore, EIS constructed according to the principles of this *Conceptual Model* in future will decrease administrative pressure, ensure on time, easy to reach, reliable and relevant information on state of environment.

REFERENCES

By-law on Environment Information System, *Official Gazette 68/08, 2008*.

Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions/*COM(2008)46 Final, Brussels, 2008.*

Croatian Environment Agency, Croatian Environment Information System Development Programme for period 2009-2012, *Zagreb, 2008.*

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Web Services and Data Harmonisation for Spatial Planning

J. Čepický[a], O. Čerba[b], K. Charvát[a], J. Ježek[b], T. Mildorf[b]

[a] *HELP SERVICE - REMOTE SENSING spol. s.r.o., Černoleská 1600, 256 01Benešov u Prahy, Czech Republic (jachym.cepicky@gmail.com, kch@bnhelp.cz)*

[b] *Katedra matematiky, Fakulta aplikovaných věd, Západočeská univerzita v Plzni, Univerzitní 22, 306 14 Plzeň, Czech Republic (ota.cerba@seznami.cz)*

Abstract: Harmonisation of geospatial spatial planning data makes possible an easier data sharing, management and publishing. This paper is focused on using and harmonisation of geospatial data in spatial planning. At first the general questions of urban planning in the Czech Republic and some aspects of data harmonisation are described. The next sections are applied to Humboldt project and its scenarios, above all Humboldt Scenario Urban Planning. This scenario represents a concrete example of a geospatial data harmonisation processes in urban planning in the Czech Republic. There are described reasons of harmonisation, harmonisations needs and requirements, current problems, proposed solutions and importance of this project and its results for other activities. This paper is based on these related projects and standards:

- Requirements of Infrastructure for Spatial Information in the European Community (INSPIRE) directive will be met, because INSPIRE aims to create the legislative and technical groundwork for the creation of a ESDI and spatial planning data should be a part of ESDI.
- Global Monitoring for Environment and Security (GMES) initiative is to enable decision makers in Europe to acquire the capacity for global as well as regional monitoring. To achieve this, GMES needs to make full use of data collected from space-borne, airborne and in-situ observation systems that is then delivered to service providers through an efficient data integration and information management capacity. Data integration is one of the fundamental tasks of GMES, therefore spatial planning data harmonisation must be in agreement with European initiative GMES.
- Reference Information Specifications for Europe (RISE) project addresses more specifically the GMES data harmonisation action line, and relates also to the INSPIRE implementation rules.
- Open Geospatial Consortium (OGC) standards – web services, e.g. Catalogue service, Web Map Service, Web Feature Service; Geography Markup Language, Sensor Observation Service etc.
- International Organization for Standardization (ISO) standards – from the view of spatial planning are very important Technical Committee 211 (Geographic information/Geomatics) standards, e.g. ISO 19115 (Geographic information – Metadata), ISO 19139 (Geographic information -- Metadata -- XML schema implementation) and others.
- Czech legislative rules.

Keywords: Spatial planning, urban planning, harmonisation processes, use case, web services, Project Humboldt.

1. INTRODUCTION – THE HETEROGENEITY AS THE FUNDAMENTAL PROBLEM OF SPATIAL PLANNING

Spatial planning represents the methods used mainly by public sector to influence the development and the use of land. But contemporary spatial planning goes beyond traditional land use planning to bring together and integrate policies for the development and use of land with other policies and programmes which influence the nature of places and how they function. Spatial planning includes all levels of land use planning including urban planning, regional planning, environmental planning and national spatial plans. Spatial planning is important not only from the view of development and stabilisation of settlement and urbanised territories. There is necessary to integrate spatial planning activities to total process of sustainable growth and landscape management. Spatial planning influences the daily life of each of us, because the spatial planning describes the localisation of public buildings, roads, recreation or industrial areas. The high-quality spatial planning makes possible the higher level of administrative decisions on the municipal policy. The publication of spatial planning activities makes possible an integration, cooperation and control of public.

Spatial planning has connection with a large number of human activities (e.g. nature protection, housing, public sector, tourism, utilities, risk management or economical activities such as agriculture, industry, transport etc.) and branches of science (e.g. geoinformatics, geomatics, geography, geology, hydrology etc.).

The widespread support and the involvement of many different subjects (commercial companies, universities, regions, municipalities, software producers etc.) evidence the importance of spatial planning, too.

The problems of spatial planning are cross-border. This cross-borderity is concerned in knowledges, terminology, legislative, methodologies, spatial data sets, software products and also the cooperation on concrete projects overlapping a territory of one state.

And just the interconnection of these above-mentioned activities, subjects and their products (software, methods, geospatial data) could reduce cost of spatial planning. Therefore the networking and other forms of cooperation are supported.

The number of involved subject means the huge number of different data sets. This heterogeneity leads to many problems associated with sharing and distribution of data, services and software products, interoperability and accessibility of final products (e.g. geospatial data sets, maps, plans etc.). The heterogeneity implicates the complicate and expensive data processing, which could result in a worse quality of final spatial plans. The background data sets used in spatial planning are not specified in the same way in different countries. On the top of it the specifications are often very general, unrealistic and uncertain. In the project Humboldt there is analysed and processed the concrete situation in the Czech Republic. In the Czech Republic there are described the source data so-called spatial (or territorially) analytic backgrounds (SAB or TAB) in the public notice No. 500/2006 (Public notice of territorially analytic background data, territorially planning documentation and methods of evidence of territorial planning activities). This public notice contains just the list of required items, which should be registered as the fundamental data sources of spatial plans. But any other specifications or properties are not contained in this legislative rule. Background data are provided by a large number of different data providers (e.g. public institutions, municipalities, large private companies, small companies etc.). Data providers hand in data in different

- data models (e.g. GIS model, CAD model, raster images, data without spatial component such as tables or texts etc.),
- data formats (e.g. SHP, GML, DGN, DXF etc.),

- mediums (e.g. web services, files on CD or DVD, printed map etc.),
- quality,
- detail,
- scale,
- portrayal rules.

This paper is focused on the summary of actual pieces of knowledge of large European projects – Humboldt (Towards the Harmonisation of Spatial Information in Europe, http://www.esdi-humboldt.eu). The experience of this project relating to geospatial data processing, especially geospatial data harmonisation could be used in spatial planning around all Europe. In addition these experience could be implemented to other activities connected with a common processing of geospatial data. In this paper there are described mainly the harmonisation processes based on web services and semantic description of these data sets and web services.

2. DATA HARMONISATION – WHAT, WHY, HOW?

In the materials of project Humboldt there is the data harmonisation defined as the possibility to combine data from heterogeneous sources into integrated, consistent and unambiguous information products.

On the present a quantity of informations is not the main problem. The quality of data is more important, because the quality results in better or worse data accessibility and usability. In our view the data quality could mean

- metadata records (data source, upgrade date, author, format, etc.),
- semantics,
- possibility of combinations with other data sets,
- implementation to used software and hardware products.

High-quality data must fulfil the above conditions without reference to data sources, state borders, scales, used technologies and platforms, data models, organization structures, legislative rules, end-user requirements, data types, data formats etc.

The need for harmonisation is due to the necessity of a cooperation of data providers, data processores and end-users on all levels. In consequence of the growing globalisation the interconnection is related to more and more economical subjects and data sets. On the present there are above all relations with EU countries very actual. Data harmonisation could have the cost reduction (e.g. costs of hardware, software, redundant data sets, manpower etc.).

Project RISE (Reference Information Specifications for Europe, www.eu-rise.org) described the parts of geospatial data heterogeneity in the year 2006. These parts of heterogeneity defined the types of harmonisation:

A. INSPIRE principles
B. Terminology
C. Reference model
D. Rules for application Schemas and feature catalogues
E. Spatial and temporal aspects
F. Multi -lingual text and cultural adaptability
G. Coordinate referencing and units model
H. Object referencing modelling
I. Data translation model/guidelines
J. Portrayal model
K. Identifier Management
L. Registers and registries
M. Metadata

N. Maintenance
O. Quality
P. Data Transformation
Q. Consistency between data
R. Multiple representations
S. Data capturing
T. Conformance (INSPIRE Drafting Team [2007])

There are three main approaches used for harmonization:

- All data sets are transformed to one common harmonized data model, before applying software or service.
- Heterogeneous data models may coexist and are transformed during runtime of software or service.
- Mix of core data model and local specialisations is used and (partial) transformation occurs during runtime.

3. PROJECT HUMBOLDT

Project Humboldt will contribute to the implementation of an ESDI (European Spatial Data Infrastructure) that integrates all the diversity of spatial data available from the multitude of European organizations, it is the aim of this project to manage and advance the implementation process of this ESDI. To achieve this objective and to maximize the benefits gained from this integration, the requirements of INSPIRE, of GMES (Global Monitoring for Environment and Security), of the environmental agencies and of other related activities in the EU will be met.
Project Humboldt started with an analysis to facilitate the re-use of existing concepts, processes, implementations and experiences. This also included the analyses of harmonization processes in other application areas. Following project Humboldt is extending the existing by the needs of users and administrators especially in the area of GMES. As a cornerstone for future businesses, spatial planning, citizen security, risk management and many more opportunities, the ESDI has to be a lasting development, prepared for the steps that will inevitably follow with the continuing progression of globalization. To enable this, the Humboldt project suggests an optimized, community-centered implementation process. New knowledge will then be gained and new processes will be developed from the possible combination of data that already exists but is currently highly scattered and heterogeneous.
Besides a technological-focussed framework which will be developed in Humboldt project, the project also will set up a number of scenarios which will use the developed framework components in real-world conditions and which will be used as promoters for the target users of the project. Several user groups like industry, public authorities and research will be targeted and therefore dissemination and training instruments are used as early as possible within the project to achieve a high coherence with the INSPIRE time plan and also to achieve high dissemination and feedback for the project results. (HUMBOLDT [2006])
The project Humboldt is devided into 12 Workpackages (WP) representing the main activities focused on the achieving of tasks of the project:

1. Administration,
2. Cost & Process Analysis,
3. State of the Art,
4. Development Process Specification,
5. Framework Interface, Models and Architecture,
6. Component Validation Process,
7. Harmonized Data Profiles,

8. Framework Development,
9. Scenario Applications,
10. Evaluation,
11. Dissemination and Exploitation,
12. Training.

An essential element of the project is the development of Humboldt scenarios. The main goal of these scenarios is employment of the developed framework components in real world conditions, based on applications that simulate these conditions identified indifferent environment-related uses cases. The scenarios are both a test bed and a community-driven research environment, which will assist in the development and promotion of the project's objectives. The following application scenarios are components of Humboldt project:

- Border Security: Enabling Effective Border Control and Security in Rural Areas,
- Urban Planning: European Urban Management Information Systems,
- Forest: Saxony & Czech Cross-Border Forest Scenario,
- Planning: Management of Protected Areas of Regional Parks,
- ERiskA: Environmental Risk,
- Water: Water Management,
- Ocean: Oil/Contaminants spill crisis impact and management,
- Galileo: GALILEO Integration for Atmospheric Data Distribution,
- Urban Modelling,
- Transboundary Catchment

(last two scenarios were adopted as midterm scenarios).

4. HARMONISATION PROCESS FROM TECHNICAL PERSPECTIVE

One of the core element of Humboldt project is set of processes that can transform the geodata to achieve the harmonization. For this propose there has been several approaches chosen depending an level of precess generality. There has been two main technologies taken into the account in Humboldt – these are:

- WPS – OGC (Open Geopsatial Consortium) Web Processing Service is gaining traction as an interface for the description and execution of transformation capabilities. It provides Web Service advantages like low coupling and is considered the preferred way to implement particular processes in Humboldt.
- WS – SOAP (Simple Object Access Protocol) – the most complex approach that should especially be taken into account when working with non-geo-specific services. However, it also becomes clear that OGC services will more commonly offer such interfaces in the future. Using this approach fits the needs of most general processing tasks for which the WPS interface is not suitable. One of the great advantage of SOAP is the amount of software tools that lets developers to easyly produce SOAP. Also BPEL (Business Process Execution Language) engines that are available for some popular IDEs (NetBeans, Eclipse) are suitable for chaining and executing SOAP services.

As far as we are thinking about complex framework that will includes not just particular WPS services but will be able also to chain them and catalogue them we have to formalized the informations that are provided about such services.
Humboldt transformation services have to provide enough metadata that describes particular WPS. Also particular metadata that will be added to resulting transformed data are necessary.
The WPS as such has to bee described using ProcessDescription. Except mandatory elements of DecribedProccess we have to use also the Metadata optional element that contains OWS:Metadata.

Humboldt WPS also has to add proper metadata to process results. The Lineage elements of ISO 19115 should be used for this propose.

5. CONCLUSION

The main goal of the project Humboldt is to enable organisations to document, publish and harmonise their geospatial data. The developed software tools, services and processes will demonstrate the feasibility and advantages of relationship of SDI (Spatial Data Infrastructure). On the basis of actual results of Humboldt project there are appearing the advantages of harmonisation of geospatial data sets in spatial planning. These advantages are very important for all data providers, data users (planners, public) and data managers – GIS department of municipalities or regions, too. Harmonisation of spatial planning geospatial data could have following benefits:

- Any duplicities in data
- Clear origin and assurance of quality of the data
- Data structure standardisation
- Data purity, security and structure uniformity
- Better data manipulation
- Reciprocal data accessing per WMS (Web Map Service) and WFS (Web Feature Service) – preservation data up-dating (possibility of on-line actualisation)
- Fall of cost for data updating and maintenance
- Better software development
- Better source exploitation
- Improvement of chances in communication with authorities
- Better utilization and commercialization of urban planning geospatial data
- Increasing activities, e.g. education

The results of Humboldt project will not be limited by

- licences (the majority of developed components will be distributed through non-commercial licences like GPL /GNU General Public License/),
- platforms (using of platform independent web services),
- territories (in Humboldt project there are used global standards, therefore the developed tools would be applicable in non-ESDI countries).

The project Humboldt and similar activities (e.g. project RISE or project MOTIIVE) try to find way from the labyrinth of data sets and different data providers – when the data are described by different methods and provided by different sources. In addition there are problems with responsible organisations, because the data provider and the organisation liable for this data are often different subjects. Therefore the spatial planning needs the semantics, metadata and uniform data models. For building of this structures there is necessary the experience sharing connected with international cooperation.

REFERENCES

HUMBOLDT Project [online]. Humboldt Consortium, 2006. URL: http://www.esdi-humboldt.eu/home.

INSPIRE Drafting Team "Data Specifications". Methodology for the development of data specifications [online]. 2007. URL: http://www.ec-gis.org/inspire/reports/ImplementingRules/inspireDataspecD2_6v2.0.pdf.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Humboldt Scenario Forest – Practical Example Forestry Data Harmonisation

J. Čepický[a], O. Čerba[b], J. Fryml[c], K. Charvát[a], M. Pospíšil[c]

[a] HELP SERVICE - REMOTE SENSING spol. s.r.o., Černoleská 1600, 256 01Benešov u Prahy, Czech Republic (jachym.cepicky@gmail.com, kch@bnhelp.cz)

[b] Katedra matematiky, Fakulta aplikovaných věd, Západočeská univerzita v Plzni, Univerzitní 22, 306 14 Plzeň, Czech Republic (ota.cerba@seznami.cz)

[c] Ústav pro hospodářskou úpravu lesů, Nábřežní 1326, 250 01 Brandýs nad Labem, Czech Republic (Fryml.Josef@uhul.cz, pospisil.martin@uhul.cz)

Abstract: The main goal of the Humboldt project is to enable organisations to document, publish and harmonise their spatial information. The proposed software tools and processes will demonstrate the feasibility and advantages of an Infrastructure for Spatial Information in Europe as planned by the INSPIRE (INfrastructure for SPatial InfoRmation in Europe) initiative, meeting the goals of Global Monitoring for Environment and Security (GMES). Finally, a software framework and diverse tools will be developed and integrated into the ESDI to support spatial data and service providers in offering standardized spatial information. Humboldt Framework is developed as open modular system. Therefore there are not any problems with updating of current harmonisation methods, adding new methods and functions, implementation to own products etc. Development and application of modern and perspective technologies (e.g. web services, expert system, knowledge databases, ontologies, semantic data sets, generalisation methods, sensor technologies etc.) could be very interested from the viewpoint of next research activities. The Humboldt project results (tools, methods, harmonisation procedures, training materials etc.) are not designed for project participants and their customers. After the end of project (2010) the Humboldt Framework and its tools will be available to all interested subjects .The Humboldt project is divided into 8 scenarios. The scenarios will implement, test and verify the results of proposed methods and developed tools. HS Forest is one of these scenarios, which is focused on forestry geospatial data in the Czech Republic. Harmonisation of forestry geospatial data could have following benefits – minimum of duplicities, application of INSPIRE recommendation, data structure standardisation, better data manipulation, reciprocal data accessing per web services and many other advantages. This paper presents the basic principals of HS Forest, processes use cases, used technologies and benefits for end-users including benefits relevant to environment.

Keywords: Spatial data, forestry, harmonisation processes, use case, Project Humboldt.

1. INTRODUCTION

The contemporary world is frequently confronted with many pressing questions dealing with environment, security, sustainable and fair growth. The informations and data are the most important tool to solution of previous questions. The informations makes possible the effective targeting of relevant precautions and decisions. On the present a quantity of informations is not the main problem. The quality of data is more important, because the quality results in better or worse data accessibility and usability. In our view the data quality could mean

- data description (metadata records, semantics, data models, ontologies etc..),
- support of system interoperability (possibility of combinations with other data sets and implementation to used software and hardware products).

High-quality data must fulfil the above-mentioned conditions without reference to data sources, data providers, state borders, scales, used technologies and platforms, data models, organization structures, legislative rules, end-user requirements, data types, data formats etc.

The need for harmonisation is due to the necessity of a cooperation of data providers, data processors and end-users on all levels. In consequence of the growing globalisation the interconnection is related to more and more economical subjects and data sets. On the present there are above all relations with EU countries very actual. Data harmonisation could have the cost reduction (e.g. costs of hardware, software, redundant data sets, manpower etc.).

This paper presents

- harmonisation of geospatial data (reasons of harmonisation, harmonisation needs and requirements, concrete questions and problems of data harmonisation, proposed solutions etc.),
- description of Project Humboldt and the basic principals of Humboldt Scenario Forest (HS Forest),
- benefits for end-users including benefits relevant to environment.

Data harmonisation in forestry is very important for following reasons:

- Forestry represents the combination of economical subject, subject of landscape management and subject of nature preservation. Forest management is also essential in term of risk management (e.g. floods or fire protection). Therefore the forestry data have the huge utilization.
- In forestry there is a large number of different data sets originated from many sources (e.g. internal data, data of state administration, data of private and state owners etc.). On that ground the data harmonisation is necessary, too.
- Above all the activities of risk management needs the very quick access to the updated forestry data (e.g. forest borders or forest roads). Just the harmonised data sets lead to better accessibility.
- The forests are situated regardless of state and other administrative borders. Therefore the support of harmonisation is very important, because the current possibilities of sharing of data sets is very poor (often impossible).

2. PROJECT HUMBOLDT

Project Humboldt will contribute to the implementation of an ESDI (European Spatial Data Infrastructure) that integrates all the diversity of spatial data available from the multitude of European organizations, it is the aim of this project to manage and advance the implementation process of this ESDI. To achieve this objective and to maximize the benefits gained from this integration, the requirements of INSPIRE, of GMES (Global Monitoring for Environment and Security), of the environmental agencies and of other related activities in the EU will be met.

Project Humboldt started with an analysis to facilitate the re-use of existing concepts, processes, implementations and experiences. This also included the analyses of harmonization processes in other application areas. Following project Humboldt is extending the existing by the needs of users and administrators especially in the area of GMES. As a cornerstone for future businesses, spatial planning, citizen security, risk management and many more opportunities, the ESDI has to be a lasting development, prepared for the steps that will inevitably follow with the continuing progression of globalization. To enable this, the Humboldt project suggests an optimized, community-centered implementation process. New knowledge will then be gained and new processes will be developed from the possible combination of data that already exists but is currently highly scattered and heterogeneous.

Besides a technological-focussed framework which will be developed in Humboldt project, the project also will set up a number of scenarios which will use the developed framework components in real-world conditions and which will be used as promoters for the target users of the project. Several user groups like industry, public authorities and research will be targeted and therefore dissemination and training instruments are used as early as possible within the project to achieve a high coherence with the INSPIRE time plan and also to achieve high dissemination and feedback for the project results. (HUMBOLDT Project [2006])

The Project Humboldt group is composed of 31 partners from 14 European countries (13 members of the EU and Switzerland). The Fraunhofer Institute for Computer Graphics (Darmstadt, Germany) is the leader of the project. On the project there are cooperated following types of institutions:

- Commercial companies (e.g. LogicaCMG, Intergraph CR spol. s r.o.),
- national mapping organizations (e.g. French National Geographic Institute, Institute of Geodesy, Cartography and Remote Sensing, Hungary),
- research institutes (e.g. Swedish Meteorological and Hydrological Insitute, Hellenic Centre for Marine Research),
- universities (e.g. Delft University of Technology, Vysoká škola báňská – Technická univerzita Ostrava, the Czech Republic).

The project Humboldt is divided into 12 Workpackages (WP). These Workpackages cover all activities of Humboldt project (e.g. planning, state-of-art analyses, development, evaluation, training etc.). The Humboldt project contain 10 scenarios (Humboldt Scenario, HS). The Scenarios are important drivers for the whole software development process and cover a wide variety of application domains, stakeholders and test areas (e.g. urban planning, hydrology, marine science, forest management, risk management, nature protection etc.).

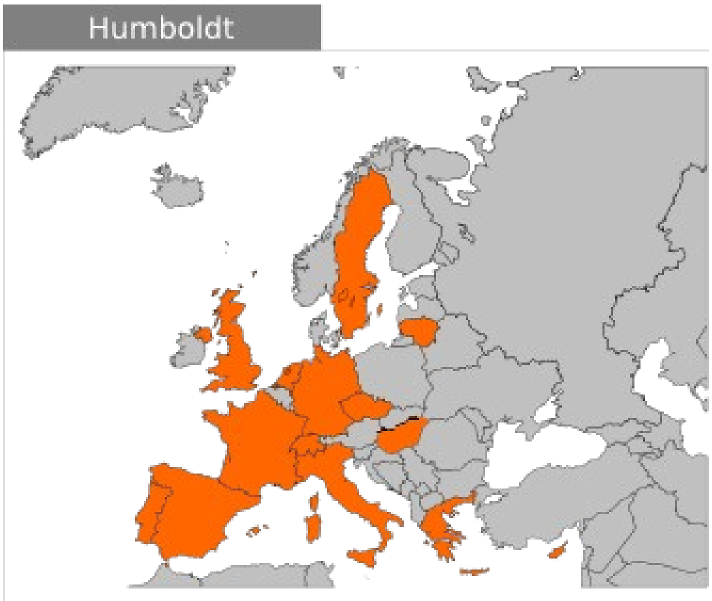


Figure 1. The countries cooperating on the Humboldt project.

3. HUMBOLDT SCENARIO FOREST

The Czech Republic, in terms of area, is one of the smaller states in Europe but it can be proud of its highly diverse natural wealth. Forests are undoubtedly one of the most valuable of its assets - their value can be expressed by purely economic indicators, but the true significance of forests far exceeds their economic importance. National Forest Programme (NFP) has been prepared in a harmony with the ongoing policy discussion on forestry issues on European level. Preparatory works on the amendment of NFP for the period 2007-2013 has already began, too. It is supposed that the main chapters of the next NFP will be dedicated to the following issues:

- Integrated forest management (including forest biodiversity in general),
- functionally differentiated forest management oriented e.g. on nature protection, protection of water and soil and social function of forest,
- forest protection,
- economic viability of forest sector,
- support of domestic processing of wood, institutes and institutions.

NFP will follow ideas of currently prepared the Forest Action Plan (FAP), which should include formulation of strategic aims particularly in following areas: biodiversity, forest protection, non-productive forest functions and benefits, usage of wood; economic viability of forest sector including non-wood goods and services.
HS Forest according NFP focuses on the management of forests in Czech Republic, concretely on metadata harmonisation, geometric and semantic harmonisation of multiple representations. There are two partners of this scenario – Forest Management Institute (FMI), Brandýs nad Laben, the Czech Republic and Help service – remote sensing, spol. s r.o., Benešov u Prahy, the Czech Republic. The development of HS Forest is focused on harmonisation of metadata, coordinate system, geometry, data models and sensor standardisation. For harmonisation tasks there is tested the Humboldt framework and its interconnection to other software products (e.g. Micka, UMN Mapserver, Intergraph

platform, Geoserver, MapMan etc.) and web services (Web Map Service, Web Feature Service, Web Coverage Service, Web Processing Service or catalogue services).

Why does FMI as the main forestry data manager in the Czech Republic need geospatial data harmonisation in forestry? There are three main reasons:

- Economical objectives: Support of a competitive strength of the forestry management, increasing of usage of forest products and services.
- Ecological objectives: Support of a biological diversity and its improvement as well as a forest resistance from local and global point of view.
- Social objectives: Increasing of life quality due to the improvement of social and cultural features of the forests.

The fundamental concrete task of harmonisation is to limit storing a managing redundant geospatial data sets. FMI offers to its clients own forestry data, but also some data (e.g. cadastral maps, contour lines form Digital Elevation Model, Geological Survey map, Soil classification map, aerial photograph, satellite image, topographic map etc.) from other providers (e.g. Agency for Nature Conservation and Landscape Protection of the Czech Republic, Information Agency for Environment of the Czech Republic, Czech Geological Survey, Czech Office for Surveying, Mapping and Cadastre etc.). These data can be bought, processed and stored in FMI, but such data will not be up-to-date at any time. The second variant is sharing data through web services, but the different data must be harmonised. The geospatial data harmonisation is very important for users of forestry data. FMI provides data to many systems of its clients. The group of target users of harmonised data contain FMI employees, forest owners, applicant for subsidies and other subjects (e.g. regional or local governments, municipalities etc.).
In terms of activities WP2 the SWOT (Strengths, Weaknesses, Opportunities, and Threats) analyse of HS Forest was processed. The main strengths of HS Forest supported harmonisation are experiences on field of digital technology in FMI (database technologies, map servers /University of Minnesota Map Server/, web services), commercial and non-commercial software experiences and detailed geospatial data description (using of metadata and catalogue system MICKA based on ISO 19XXX standards). This analyse shows also weaknesses and threats of harmonisation processes in FMI (e.g. data and platform heterogenity and financial limits), but from our point of view they could not endanger the target of Humboldt project. SWOT analyse offers for scenario partners and for forestry in general terms many opportunities (e.g. clear origin and assurance of quality of the data, data structure standardisation, forest community support etc.), which could be very interesting for many reasons (economical, social etc.).

4. HARMONISATION PROCESSES

The concrete harmonisation processes in HS Forest are focused on following attributes or properties of forestry data:

- Data format – the target data format would be to achieve syntactical interoperability by establishing data exchange via a standardised metadata catalogue (including a Feature Catalogue), and web services (WMS, WFS, WPS, WCS, etc).
- Spatial reference system – the spatial reference systems differ between Czech Republic, Poland, Germany and Slovakia. The current solution is depending on the situation. Coordinate transformations to another system are usually processed directly in the knowledge database. For the use cases described in the Scenario coordinate transformation by a WPS or WCTS would be useful.

- Conceptual data model – the conceptual data models for the basic geographic reference information differ between the Czech Republic, Poland, Germany and Slovakia.
- Classification – it is significant to have standard classifications to follow the development of landscapes from the point of view of forest damage, biodiversity, erosion, and conservation policies. For all attributes having a list of possible values (an enumeration attribute) the Scenario will try to use already existing standard (official or not official) classifications, e.g. for the type of forest.
- Terminology - the terminology used in the forest management field should be published in a (multi-lingual) thesaurus or ontology.
- Metadata – metadata records are very important for this Scenario. There is a Metadata portal for Czech spatial datasets, called MICKA, that is able to access different kinds / versions / profiles of Catalogue Services. In the context of the discussion about the core HUMBOLDT metadata profile, the Scenario will list the metadata elements that should be mandatory from the perspective of the Forest Scenario.
- Portrayal rules – a possible solution could be to define presentation styles (via Styled Layer Descriptors, SLD) and to connect them with context information. The context profile must be developed for each user that also contains information about prefered portrayal solutions.
- Processing functions – harmonisation of some processing functions should be looked at, e.g satellite imagery classification, tolerance values etc. in the calculation of spatial intersections for the subsidy calculation, defining of the conflict areas between the commercial interests in the forest management and the principles of nature conservation.
- Multi-linguality – the end-users will use the system in their own language, therefore it is necessary to create lingual transformations.
- Spatial consistency – in this Scenario there is data harmonisation not only from the data model point of view, but also from the geometry point of view: the same real-world objects exist in a different layers and map compositions, and this must be dealt with (removing or merging redundant items).

The proposed Humboldt solution of geospatial data harmonisation is based on these related international projects and standards (Čerba et al. [2008]):

- To achieve the objectives of Humboldt project and to maximize the benefits gained from the integration, the requirements of Infrastructure for Spatial Information in the European Community (INSPIRE) directive will be met, because INSPIRE aims to create the legislative and technical groundwork for the creation of a ESDI and Humboldt solution should be a part of ESDI. Document of INSPIRE Drafting Team "Data Specifications" called Methodology for the development of data specifications define data harmonisation components used in Humboldt project.
- Global Monitoring for Environment and Security (GMES) initiative is to enable decision makers in Europe to acquire the capacity for global as well as regional monitoring. To achieve this, GMES needs to make full use of data collected from space-borne, airborne and in-situ observation systems that is then delivered to service providers through an efficient data integration and information management capacity. Data integration is one of the fundamental tasks of GMES, therefore Humboldt project and its results must be in agreement with European initiative GMES.
- Reference Information Specifications for Europe (RISE) project addresses more specifically the GMES data harmonisation action line, and relates also to the INSPIRE implementation rules. Humboldt project uses for example the

harmonisation requirements questionnaire from RISE Methodology and Guidelines on Use Case and Schema Development.
- Open Geospatial Consortium (OGC) standards – in the Humboldt project there are used standards defined web services, e.g. Catalogue service, Web Map Service, Web Feature Service, Geography Markup Language, Sensor Observation Service etc.
- International Organization for Standardization (ISO) standards – from the view of Humboldt Scenario Forest are very important Technical Committee 211 (Geographic information/Geomatics) standards, e.g. ISO 19115 (Geographic information -- Metadata), ISO 19110 (Geographic information -- Methodology for feature cataloguing), ISO 19139 (Geographic information -- Metadata -- XML schema implementation) and others.
- Czech legislative rules and internal regulations of FMI are also the parts of Humboldt Scenario Forest.

5. CONCLUSION

Forests play important roles in the protection of soil or the surface under the forest cover, for instance, for protection against erosion. Forests are also essential for the maintenance of water resources and of water cycles such as the protection of water reservoirs or filtering of water, modification of water cycle and run-off. In addition, protective forests guarantee other important ecosystem functions, like the maintenance of clean air, stabilization of local climate, securing the timber line in alpine etc.
That is why the need of quality informations for the decision making on the forestry management is very important. The extending possibilities of all-round utilization of geospatial information in various spheres of human activities, above all in decision-making processes, generate preconditions for economic development, improvement of public administration and support of sustainable environment.

The main goal of Humboldt project is make possible to subjects processing the spatial data documentation, describing, publishing and sharing their spatial data sources (Čerba et al. [2007]). That's why the project was named after famous German scientist Alexander von Humboldt (1769-1859), because this eminent personality of European science proposed integrating of knowledges across Europe.

After the end of project the Humboldt framework (concrete tools, web services and methods) will be at the disposal. During the two years of the project there were finished some following activities:
- Successful Review Board meetings,
- cost analyses,
- HUMBOLDT brochure,
- State of art in spatial data tools (Humboldt document A3.2-D1),
- Handbook of Standards (Humboldt document A3.6-D1),
- Development of Humboldt Tools (e.g. Humboldt editor),
- Review of licences of geoinformation technologies,
- Dissemination activities (e.g. INSPIRE conference 2008, IST Africa 2008, GIS... Ostrava 2009 etc.).

Many other activities (e.g. creation of data models, metadata profiles, harmonisation requirements, evaluation requirements or processes specification) are under development.

REFERENCES

Čerba, O., Fryml, J., Pospíšil, M., Podlena, R., Charvát, K. HS Forest Scenario of Humboldt - Way to INSPIRE Implementation in Forestry. In Informační systémy v zemědělství a lesnictví. Praha: Česká zemědělská univerzita v Praze, 2007.

Čerba, O., Fryml, J., Charvát, K., and Pospíšil, M. Humboldt for Forestry Data Harmonisation. In IST-Africa 2008 Conference Proceedings, Paul Cunningham and Miriam Cunningham (Eds), IIMC International Information Management Corporation, 2008, ISBN: 978-1-905824-07-6.

HUMBOLDT Project [online]. Humboldt Consortium, 2006. URL: http://www.esdi-humboldt.eu/home.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

SAS for Sustainability Management: Enabling Green Innovation

A. Farrell[a]

[a] *SAS Global Headquarters, Cary, North Carolina, United States (alyssa.farrell@sas.com)*

Abstract: Based on decades of work in business analytics applied to environmental performance, SAS will review a framework for addressing sustainable innovation with information technology. This paper will provide perspectives on the role that information technology plays in addressing environmental performance management, and comment on emerging trends, challenges and opportunities.

Keywords: Sustainability, ICT, Modelling, Measurement, Analytics

1. INTRODUCTION

All organizations are called to become better stewards of our natural resources, embrace environmentally friendly practices, adopt fair trade policies, and the urgency to implement "sustainability strategies" has shifted from a small but vocal number of advocates to broad majorities of consumers and governments worldwide. Issues such as climate change, energy consumption, labor practices, food safety, pollution, and waste management are strong factors in the impressions that organizations make with stakeholders such as consumers, investors, regulators, and watchdogs.

In late 2008, a study by the Economist Intelligence Unit explored the hypothesis that corporate citizenship can help to improve the bottom line. Seventy-four percent of respondents to the survey conducted for this report say that corporate citizenship can help to increase profits at their company. In an economic downturn, it can be a vital competitive advantage.

The same can be true for government organizations who seek to be more efficient with their use of energy, water, and waste. By improving resource use, they can reduce costs. What is required, however, is a holistic picture of environmental performance, areas of concern and opportunity, and the modelling capabilities to determine the correct policy framework.

Regardless of industry, size or location, strategic management is fundamentally about identifying opportunities and obstacles and navigating through the external environment in a way that makes the most of the organization's assets. The key, of course, is to invest the right amount of time and manpower to carefully research and understand all sustainability-affecting aspects of the business. Organizations must find the right data to describe and measure those metrics. Predictive analytics tools can help identify the proper leading indicators that predict and measure sustainability performance.

2. A HOLISTIC SYSTEMS-BASED APPROACH FOR ENVIRONMENTAL MANAGEMENT

In *The Necessary Revolution: How Individuals and Organisations Are Working Together to Create a Sustainable World*, Peter Senge of the Sloan School of Management, MIT reveals that companies will need a deep systems-based understanding of how the global economy, environment, society and geopolitics interact and affect the organisation. His work predates the current economic crisis, but it only strengthens his argument. Add in volatility in energy and natural resources markets and potential disruptions in resource supply, and the importance of large-scale system-based comprehension becomes crucial for companies to succeed or, in some cases, simply survive [Vermeer, 2009].

As governments, businesses and the not-for-profit sector respond to the critical imperative of environmental responsibility – and discern its social and strategic importance – they are learning the vital steps of integrating and analysing data to achieve new goals and transform their internal culture.

Today in nearly every industry, organizational performance often integrates sustainability measures across social, environmental and economic factors. The challenge is learning where to start: deciding which initiatives will have the greatest positive impact on achieving stated goals and how to turn complex information into effective, cost-conscious strategies. Forward-thinking organizations, both private and public, are formalizing their data tracking, accounting and reporting processes using business intelligence, scorecards and advanced analytics.

3. FOUR STEPS TO RESPONDING TO THE NEW MANDATE

What is measured really matters – and technology can help improve the business' response to the sustainability mandate in four ways:

MEASURE – By properly addressing the need to integrate and validate data, organizations can benchmark key sustainability activities using industry-accepted methodologies and protocols – such as those promulgated by the EU, World Resources Institute (WRI), and World Business Council on Sustainable Development (WBCSD). Ultimately, the greatest challenge is accessing and trusting the data.

REPORT – As with economic performance, it's essential to fully and formally disclose sustainability performance to ensure transparency and alignment with key stakeholders. By regularly disclosing an integrated, consistent source of quality information, organizations can bind initiatives to a common sustainability framework that promotes consistency across multiple geographies.

IMPROVE – After identifying and measuring the metrics that have the greatest impact on sustainability goals, policy makers can make more informed strategic decisions. Applying optimization, forecasting and data-mining capabilities to analyse scenarios and run simulations improves successful policy design and implementation.

FORECAST – Finally, a solid foundation of aggregated and well-structured sustainability data enables multiple public sector organizations to manage the resources needed to achieve the desired environmental outcomes. Key strategies here include aligning investments in renewable energy, building efficiency, air and water quality, economic development, and market-driven emissions management.

4. HOW ICT CONTRIBUTES TO INTELLIGENT ENVIRONMENTAL MANAGEMENT

SAS is working with leading international organizations to apply technology to the global issues of greenhouse gas emissions, as well as other environmental and social concerns. We understand the need to deliver reliable measurement of performance indicators, as well as the predictive ability to validate strategies and costs, identify causal relationships, and forecast outcomes.

4.1 US. Environmental Protection Agency in partnership with Environment Canada

Few things are as important as the air we breathe. And since no one is immune to the effects of air pollution, it's important to understand air quality levels around the world. The Air Quality Research Branch (AQRB) of Environment Canada and the US Environmental Protection Agency count on SAS to help measure and analyse trace gases in the atmosphere and to determine the impact of human activities on air quality [SAS, 2005].

Challenge:
Manage, analyze and share air quality data from more than 40 sites.

In 1991, Canada and the United States signed an Air Quality Agreement to address transboundary pollution, which includes commitments to reduce the major pollutants that cause acid rain – sulphur dioxide and nitrogen oxides. Both countries are cooperating to assess the impacts of air pollution and transboundary pollution on human health, forest ecosystems and surface waters. The right shared information system is important to these cooperative efforts in data analysis, modelling, monitoring and information sharing.

When they began using SAS at AQRB, data volumes were increasing at a tremendous rate, and SAS provided a scalable solution to manipulate, visualize and analyse the data. In addition, SAS has been used to develop a data-sharing standard that makes it easy to exchange data with these agencies for joint research projects and international studies.

4.2 Energy Efficiency at Poste Italiane Group

In 2004, Poste Italiane Group made the courageous decision to start Progetto Speciale Energia (the Special Energy Project), headed by expert Energy Manager, Luciano Blasi, and with the support of the organization's IT department. The goal was simply defined: monitor and optimize electric energy, fuel and water consumption at all government real estate properties.

"My main purpose is to reach our energy efficiency goals in at least 250 facilities," Blasi explains, "including those with the highest energy consumption, such as data processing centres, executive centres and perhaps the largest branches. Of course, reducing costs is a goal, but that is preceded by goals for awareness, consumption control, sustainability and reducing our environmental impact [Mariani, 2009]."

"We are dealing with an overall structure that includes 14,500 buildings, half a terawatt per year in electric power, 8.5 million litres of heating oil and other heating fuels, and gas consumption of more than 20 million cubic meters."
Luciano Blasi
Energy Manager, Poste Italiane

Before beginning this project, the IT structure of Poste Italiane was focused on individual productivity platform with fully integrated databases. At the time, their purpose was mainly management control rather than energy monitoring. The billing statement served as the starting point for a preliminary analysis of consumption and subsequent monitoring, as it contained simple data elements that wouldn't require a complex, shared platform.

Before 2006, the calculations were made parametrically, based on occupied square footage. However, a more accurate analysis of consumption based on actual meter readings can identify peaks or anomalies – and thus points of intervention – by informing the facility manager, who directly reviews energy costs.

A further complicating factor is the large number of buildings in operation, which makes it impossible for Poste Italiane to use a traditional, off-the-shelf energy audit tool for analysis, demand management and efficiency improvement.

Therefore, Post Italiane needed a business intelligence tool that would process only the acquired data and could produce forecasts based solely on this data.

With the right ICT from SAS, they can move from using billing data to accurately collected measurement data. Therefore, it will be possible to model alternative strategies, know when to change policies, and correct operation and maintenance behaviour for systems.

Currently, the system is a virtual distributed structure where data pertaining to the individual locations is stored and analyzed. But the future intent is to allow each territory manager to directly input and receive information from SAS and monitor his own consumption by means of a Web-based tool.

This is the second stage of the SFINGE (Sistema di Flussi Integrati per la Gestione Energetica or Energy Management Integrated Flow System) project, and Blasi expects significant developments in various areas. For example, the process of instilling responsibility in the facility managers will lead to the definition of energy performance objectives, and hopefully the appointment of energy managers within each territory. Blasi says he will also continue with the policy of taking small steps that can add up to real energy savings.

5. CONCLUSIONS

Whether modeling cleaner generation of power, improving energy efficiency, promoting better utilization of critical resources to minimize waste, or assisting environmental protection agencies, it is imperative to have a holistic picture of environmental performance. Establishing an integrated system has benefited both public and private institutions by forecasting future emissions, identifying areas of potential reduction, and guiding new policies and procedures that will benefit global communities for future generations.

REFERENCES

Mariani, Valerio. Poste Italiane Group: A handful of data, a lot of energy savings. *Computer Business Review Italy*. http://www.sas.com/success/poste.html

SAS, Environment Canada Breathes Easier. 2005. www.sas.com/success/environmentcanada

Vermeer, Daniel and Robert Clemen, Why Sustainability is still going strong. *Financial Times,* February 12 2009.

The European Biodiversity Observation Network - EBONE

Ľ. Halada[a], R.H.G. Jongman[b], F. Gerard[c], L. Whittaker[d], R.H.G. Bunce[b], B. Bauch[e] and D.S. Schmeller[f]

[a] *Institute of Landscape Ecology of Slovak Academy of Sciences, branch Nitra, Akademicka 2, 949 01 Nitra, Slovak Republic (lubos.halada@savba.sk)*
[b] *Alterra, Wageningen UR, The Netherlands (Rob.Jongman@wur.nl; bob.bunce@wur.nl)*
[c] *Centre for Ecology and Hydrology,Wallingford, UK(ffg@ceh.ac.uk)*
[d] *Israel Nature and Parks Authority, Jerusalem, Israel (linda.whittaker@npa.org.il)*
[e] *UFZ-Helmholtz Centre for Environmental Research, Leipzig, Germany (bianca.bauch@ufz.de)*
[f] *Station d'Ecologie Experimentale du CNRS à Moulis, St. Girons, France (dirk.schmeller@EcoEx-Moulis.cnrs.fr)*

Abstract: EBONE (European Biodiversity Observation Network) is a project developing a system of biodiversity observation at regional, national and European levels as a contribution to European reporting on biodiversity. The project focuses on GEO (Group of Earth Observations) task BI 07-01 to unify many of the disparate biodiversity observation systems and creates a platform to integrate biodiversity data with other types of information. The system will make use of existing networks of site observations, wider countryside mapping and Earth Observation (EO). The project addresses issues important for development of biodiversity monitoring system e.g., the underlying concepts, selection of indicator species and habitats, in-situ and EO methods of biodiversity; database management and IT tools; protocols and harmonisation of available in-situ data. Special attention is paid to inter-calibration of in-situ and EO monitoring. The system, methods and protocols developed in the project will be tested and validated in the field. Refinements to the system, involving sites and protocols, will be proposed, based on this validation. The project aims to contribute to a world-wide monitoring system by developing a prototype for Mediterranean and related desert biomes outside Europe. Because of this wide scope, stakeholders will be involved in the design, development and testing of the monitoring system. The main outcome will be an integrated monitoring system based on key biodiversity indicators which will be implemented within an institutional framework at the European level.

Keywords: Biodiversity observation, GEOSS, GEO-BON, FP7 project, Monitoring harmonization, Information system

1. INTRODUCTION

Measuring and reporting reliably trends and changes in biodiversity requires that data and indicators are collected and analysed using standardized procedures. This is valid for a national park, but also for larger areas such as the European Union. However, at present the responsible authorities including over 100 national and regional authorities have different and uncoordinated approaches. Globally, the problem is even more serious, as the variability of species and ecosystems is greater in the different continents. There is therefore an urgent need to develop a coherent procedure for data collection and analysis that can be used for international comparative assessments of biodiversity and monitoring.

EBONE is a Collaborative project of the 7th Framework Programme within theme 6 Environment (Topic 4.1.1.2. Contribution to a Global Biodiversity Observation System). The EBONE project is developing a system of biodiversity observation at regional, national and European levels as a contribution to European reporting on biodiversity as well as to the Global Environmental Observation System of Systems (GEOSS) tasks on biodiversity and ecosystems. EBONE assesses existing approaches on validity and applicability initially in Europe, but expanding to regions in Africa and seeking cooperation with projects in other continents. EBONE aims to deliver:

1. A sound scientific basis for the production of statistical estimates of stock and change of key indicators that can then be interpreted by policy makers responding to EU Directives regarding threatened ecosystems and species;
2. The development of a system for estimating past changes linked to the forecasting and testing of policy options and management strategies for threatened ecosystems and species.
3. A proposal for a cost-effective biodiversity monitoring system in close collaboration with the major agencies and Non Governmental Organisations (NGOs) that will be responsible for monitoring in the future.

This paper therefore deals with both technical and with organisational aspects.

2. EXISTING MONITORING PROGRAMS IN EUROPE

The information on running monitoring programs in Europe is based on an ongoing survey of biodiversity monitoring schemes in Europe, started by the EuMon project [EuMon 2009; see also http://eumon.ckff.si/) and continued in the EBONE Project. The first results of that survey demonstrate that biodiversity monitoring is lacking a standardized approach in Europe, making it difficult to assess the state and trends in biodiversity across geographical and temporal scales [Schmeller 2008]. The EuMon survey is summarized in the online database DaEuMon [EuMon 2009] and allows to draw a detailed picture of monitoring practices in European species and habitat monitoring [Henry et al. 2008; Kull et al. 2008; Lengyel et al. 2008; Schmeller 2008; Schmeller et al. 2008]. Biodiversity monitoring in Europe encompasses a wide diversity of different habitat and species monitoring programs. Which vary in sampling effort, methodology (Figure 1), involvement of volunteers, incentives, and geographical scope. 50% of all habitat monitoring schemes (here a monitoring program may encompass various monitoring schemes for different species or habitats) are at local scales, followed by regional (28%) and national (18%) scale. For species it is the opposite, with schemes on a national scale making up 42% of all schemes followed by regional (27%) and local (25%) schemes. Only 0.6% (habitats) or 1.6% (species) and 2.4% (habitats) or 2.7% (species) of the assessed monitoring schemes cover the EU or are of international scale, respectively. Generally monitoring schemes were implemented due to scientific interest (31%), EU directives (21%), management or restoration (20%) or national obligations (17%). The support of volunteers is generally higher in species schemes (86% of all persons involved, see also [Schmeller 2008] and references cited therein) than in habitat schemes (9% of all persons involved) [Eumon 2009]. For habitats and species sampling effort varies greatly across spatial and temporal scales [Lengyel et al. 2008; Schmeller et al. 2008; EuMon 2009].

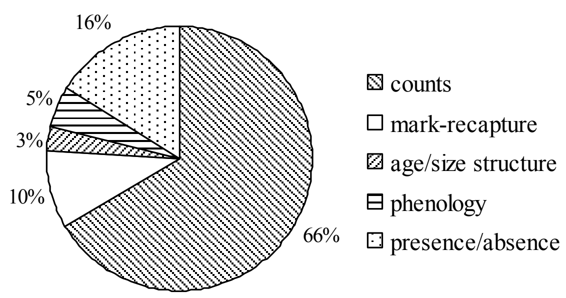


Figure 1. Monitoring methods used in species schemes. Source: EuMon Database on European Monitoring Schemes, February 2009, http://eumon.ckff.si/

According to the EuMon database, birds are the most widely represented species group in monitoring programs followed by mammals and invertebrates (Figure 2). Monitoring of birds is mainly coordinated by BirdLife International and its partner organizations within the European countries [BirdLife International 2004]. In parallel to the birds scheme, butterfly monitoring in Europe [van Swaay et al. 2008] also provides a good blueprint for the establishment of monitoring networks for other groups across Europe [Gregory et al. 2005; European Environment Agency 2007]. Butterfly monitoring schemes assess regional and national trends in butterfly abundance per species, and form a network for the coordination of data in different countries [van Swaay et al. 2008].
For European habitats and species the EuMon survey demonstrates that monitoring needs to improve with regard to spatial coverage, assessment of spatial variation, sampling design and data analysis. Without these improvements monitoring data will not be able to present an unbiased and realistic picture of the state of Europe's biodiversity that is necessary to measure the progress towards halting the loss of biodiversity by 2010 [Lengyel et al. 2008].

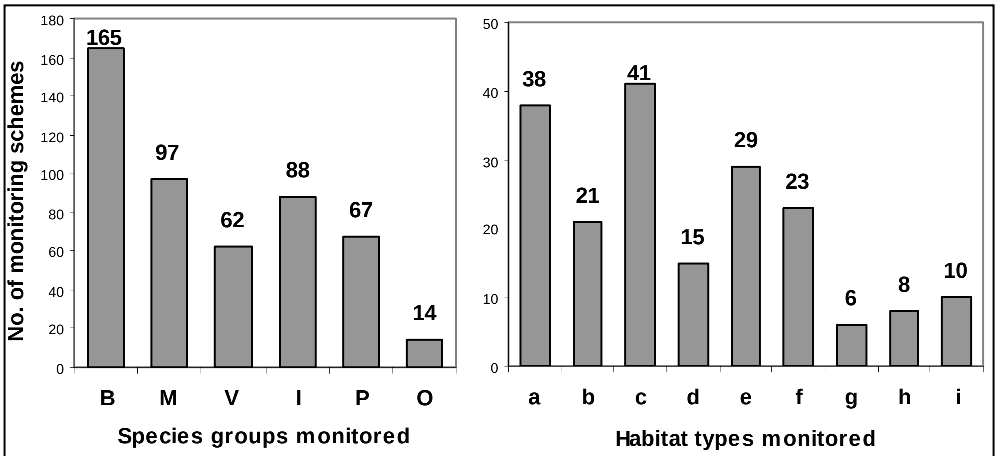


Figure 2. Number of monitoring schemes per species group and per habitat types in Europe
B – birds; M – Mammals; V – other Vertebrates; I – Invertebrates; P – plants; O – other; a - coastal and halophytic habitats; b - coastal sand dunes and inland dunes; c - forests; d - freshwater habitats; e - natural and semi-natural grasslands; f - raised bogs, mires and fens; g - rocky habitats and caves; h - sclerophyllous scrub; i - temperate heath and scrub. (Source: EuMon Database on European Monitoring Schemes, February 2009, http://eumon.ckff.si)

However, some of these improvements have already been implemented in three major international projects which are not included in the EuMon database:

1. The Countryside Survey of the United Kingdom [Haines-Young et al, 2000]. This survey has tracked habitat and vegetation changes from 1978 to 1984, 1990, 1998 and most recently 2007.The sample has increased from 256 1 km squares to over 600 at present. An extensive literature reports the results.
2. SISPARES [Ortega et al. 2008] is a series of over 200,16 square km units which have had aerial photograph interpretation over three dates since the 1960's to enable the assessment of habitat change in Spain.
3. The National Land Inventory of Sweden [Esseen et al. 2006] consists of a series of sites throughout Sweden with the objective of assessing habitats and vegetation for the whole country and in due course, the monitoring of change.

All these projects are at a national level and are based on stratified random sampling to ensure that statistical estimates of habitat extent and spatial distribution can be achieved.

3. EBONE PROJECT

3. 1 EBONE concept

Despite the wide range of monitoring activities described in the EuMon database, existing monitoring schemes cover only specific elements of biodiversity and usually have a restricted geographical coverage. They also often have limitations regarding the potential for spatial interpolation and generalisation as well often not covering long time periods. The limited data that have currently been collected systematically can only enable restricted statistically reliable conclusions to be drawn. However, for Europe-wide monitoring and reporting the incorporation of further data are essential to extend the range of systems to be covered.

The project is designed to respond to the widely recognised problem of limitations in the linkage among existing monitoring systems, databases and monitoring sites. The key challenge that needs to be addressed is the cost effective data collection system for biodiversity linked with extant data, both past and present, at national, regional and European levels. This monitoring framework should cover defined aspects of biodiversity in one coherent system. A systematic monitoring approach for Europe must consist of several steps and every action for collection of new data will first need to consider what existing data are available and how they can be used and interpreted. Because cost-effectiveness is a crucial consideration in the design of monitoring systems, an assessment of the time and costs involved will also be carried out.

It is essential that the scientific basis is linked to a sound institutional framework to ensure continuity and long-term collaboration between partners who have a history of successful cooperation. The project includes facilitation of an institutional cooperation in a stratification system, nomenclature and data to be collected and agreements on database structure. The present project is based on a tried and tested network of partner institutes that have collaborated over many years and that have been monitoring change at a variety of scales. The network is also open to other interested partners who may wish to join.

The strength of the approach is that it builds on the knowledge and networks developed in recent European projects such as ALTERNET, BioHab, BioPress, BioScore, Ecochange and EuMon. It makes use of existing LTER monitoring sites; it assesses their representativeness and includes the existing national monitoring systems. It will lead to a cost effective procedure for biodiversity monitoring by applying the most efficient indicators in a well balanced sampling programme.

3.2 EBONE implementation

The first phase of the project is to develop a conceptual framework for monitoring, the utilisation of the existing institutional context of European monitoring, databases, observation points and observing organizations agencies, and NGOs. The criteria for identifying indicators will be defined using existing experience and the framework of the CBD and SEBI and going beyond if needed. The design of requirements, protocols and procedures for a cost-effective monitoring system for Europe requires bringing together existing knowledge on monitoring protocols and a concept that is able to upscale and downscale data and observations from point locations to a general European level. It also needs a concept of the sampling design that can be used to test the existing data, observation points and databases. The conceptual framework of the sampling design is also required to test the existing data, observation points and databases. It will also be used to consider how monitoring of biodiversity trends can be linked with site specific ecosystem research on underlying processes, drivers and pressures.

An overview will be prepared of the characteristics of the existing larger monitoring and surveillance systems in Europe. This will then be used to provide links between the methods, data and observation sites available in different countries and regions as well as with various ongoing projects, available databases and observation and monitoring systems. The relationship between National Responsibilities (NR) and Conservation Priorities (CP) of species and habitats with monitoring systems will also be determined. A proposal will be made for the best way to integrate different monitoring systems into a European strategy that includes NRs and CPs. The experience from this work will be used to develop an approach for a global biodiversity observation system.

A statistically robust framework for monitoring is under development and it will form the basis for a European system for reliable, geographically referenced and comparable data collection of species and habitats of conservation interest. The existing knowledge on monitoring procedures will be used for harmonization of protocols for different species groups and habitats. The General Habitat Categories of the BioHab project will be used as a common denominator to link existing data sets. The special attention will be paid to inter-calibration of Earth Observation (EO) and in-situ monitoring data (see chapter 4). The monitoring system will be validated and the cost aspects in time and budget will be checked in representative test sites. The sample sites in the project will be dispersed in strata defined in the project of the European Environmental Stratification [Jongman et al 2006]. One of the important steps is to carry out tests on the data from LTER (Long-Term Ecological Research) sites in relation with data from nation-wide habitat monitoring programmes. The system developed for Europe will also be adapted for Mediterranean systems in test areas in Israel and South Africa as representative countries for this biome (see chapter 5.3). The institutional arrangements and the cost effectiveness of the proposed surveillance and monitoring methodology will be evaluated and the results used to design a management procedure for a time- and cost efficient monitoring system.

Data management represents an important aspect of the monitoring system developed in the project; the main challenge being to provide a common database access across all types of data supplied by diverse sources and in varying granularity, ranging from remote sensing data over national monitoring networks to the very dense and detailed data available from the LTER sites.and LTSER platforms.

In all stages of the project stakeholder involvement is required; important is the communication to agencies institutions, managers of databases and organisations that already carry out established European-wide and countrywide monitoring systems. These stakeholders will be involved in the development of the monitoring system, its testing and improvement as well as possible instruction in its application The communication with international, national and regional responsible bodies is essential to get the system accepted and agreed upon.

4. EO OBSERVATION AND INTER-CALIBRATION WITH IN-SITU MONITORING

4.1 EO observation of land cover, habitats and species

EO instruments record reflected, scattered or emitted electromagnetic signals which vary in function of the physical and chemical properties of the viewed surface type. Two types of information can be derived from EO imagery (Figure3): quantitative measures of these physical or chemical properties (i.e. a map of for example soil moisture, surface temperature or canopy cover) or a map of thematic classes representing areas with similar reflected, scattered or emitted electromagnetic signals, texture, patterns or shapes. EO derived products of land cover, habitats and species (flora) belong to the second category.

The observation and recording of land cover, habitats and species require different classification systems. Their design results from a compromise between scope of use, level of detail and spatial application. EO introduces not only full area and frequent coverage, but also a new and unique set of classification parameters. The degree in which a relationship can be established between electromagnetic signals and the thematic classes (e.g. physiognomic, floristic or ecological) required by the biodiversity monitoring community, will determine the usefulness of the EO derived thematic maps. The work of Paradella et al. [1994] suggested that physiognomy may be the most important attribute which influences the EO response of vegetation. Whilst Jakubauskas et al. [2002], Moody and Johnson [2001] and Hill et al. [submitted] have reported successful crop, vegetation and species classifications when using time series of EO to exploit differences in phenology. Many have shown that when working regionally or locally, and using EO data types and classification approaches appropriate for the local scenario, accurate and reliable and therefore useful results can be achieved [Hill and Thomson 2005, Thomson et al 2003, Bock et al 2005]. However, when continental and global biodiversity monitoring requires consistency in methodology, the variety of EO data types and approaches available is greatly reduced. As a result, the global land cover maps produced from EO have been limited to reporting the extent of major vegetation types (total number of vegetation classes ranges between 7 and 18, Table 1) at pixel sizes ranging from 1km to 300m. The class number and type and the spatial detail of these products make them inadequate for detailed biodiversity or habitat monitoring.

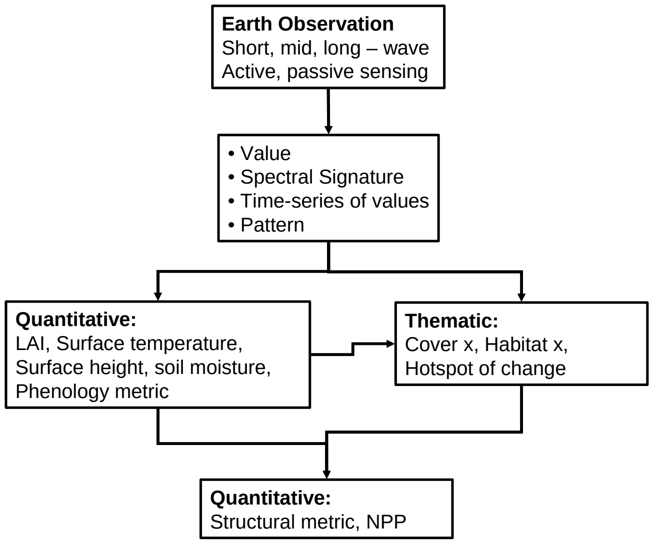


Figure 3. Schematic illustrating the difference between the quantitative and thematic measures derived from Earth Observation.

In addition to thematic maps, EO can deliver quantitative information that is related to site conditions, physiological processes, stress conditions or vegetation damage, and is relevant to biodiversity. For example, the leaf phenological cycle and its changes over time have been measured with EO [Delbart et al. 2006, Heumann et al. 2007], the SEBI indicator 'fragmentation' is an obvious candidate for EO retrieval [Estreguil et al., submitted], and EO vegetation indices have been related to NPP and linked to species richness [Oindo and Skidmore 2002].

Table 1: Global Land cover maps derived from EO currently available

Land cover map	Pixel size	Total N^o classes	N^o vegetation (arable) classes
IGBP [Loveland and Belward, 1997]	1 km	17	12(2)
GLC2000 [Bartholome and Belward, 2005]	1 km	22	18(3)
MOD12Q1 PFT [Friedl et al., 2002]	1km	11	7(2)
GLOBCOVER [Arino et al., 2005]	300 m	22	14 (4)

'Going in situ' is the only way to collect detailed information on the flora and fauna present. Also in situ land cover or habitat observations, when benefiting from a well designed field survey approach and protocol, have the advantage of providing high thematic and spatial detail. In both cases, in situ work is intensive and costly and is therefore limited in the area it can cover and the revisit frequency, and although the spatial detail of the area outlines identified can be high, they often vary with surveyor, especially in areas containing soft gradients rather than hard boundaries.

4.2 Inter-calibration of EO and in-situ monitoring

In the context of biodiversity monitoring, the idea of integrating in situ with EO is that the combination of the two data set types will deliver more accurate or reliable information on biodiversity than either of the two data sets used independently. EO can take on different roles when it is considered for enhancing in situ observations. It can be used as a vehicle for interpolation and generalisation by delivering full coverage, or it could be used to increase the number of in situ samples in space and time. Here, the key for success is a good link between the EO derived thematic map and the in situ habitat observations. EO could also be used to search for and highlight hotspot areas of sudden or gradual change, or to provide context where it delivers additional information on, for example, land cover composition, landscape structure or phenology, complementing the in situ species and habitat data. In all cases the concept of linking EO derived information with field data to enhance observations on biodiversity is based on the premises that a relationship exists between the composition and structure of the landscape and the diversity of habitats and the species and genotypes that may be present within.

Many trials have already been carried out, and although much discussed, full integration between in situ and EO has not been achieved, as emphasised in the recent GMES summary produced by Wyatt et al [2004]. The BIOPRESS project [Köhler et al 2006] and the PEENHAB project [Mücher et al. 2004] whilst using the state of the art data bases for predicting habitats, show that all the available data bases have limitations and restrictions because of lack of validation. EBONE is planning to provide clear statements on the added value of data integration by testing if integration delivers improved estimates of biodiversity measures, in particular the SEBI indicators: (i) Trends in extent of selected ecosystems and habitats and (ii) Trends in abundance and distribution of selected species.

One main element of the work will investigate an approach developed by Fuller et al [1998] for the UK which is referred to as 'inter-calibration'. Inter-calibration is particularly suited for the SEBI indicator 'Trends in extent of habitats'. Inter-calibration uses correspondence matrices [Lillesand and Kiefer, 1994] that are created to calculate the classification accuracy of EO derived land cover maps, to produce stratum specific calibration matrices (i.e the calibration matrices are a weighted average of the correspondence matrices produced for in situ sample sites located within an environmental stratum). These calibration matrices are then used to alter the original land cover percentages and classes of the EO derived map to better match the habitat classes and their cover percentages observed in the field (Figure 4). Although this reduces the original spatial resolution of the land cover map from 25 m to 1 km, Fuller et al [1998] found that,

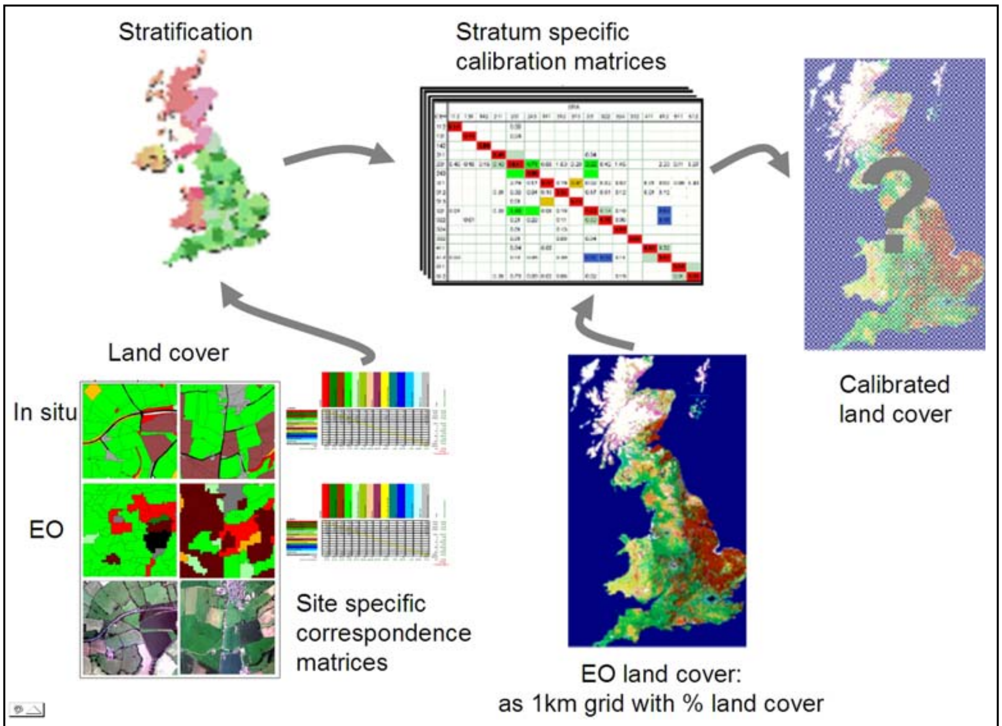


Figure 4. Diagram illustrating inter-calibration as implemented by Fuller et al (1998)

at national level, the habitat statistics produced from the calibrated land cover map closely matched those extrapolated from the field samples. More importantly regional habitat estimates made from the calibrated land cover map showed tighter confidence limits than those acquired from the field survey samples.

The EBONE hypothesis is that better estimates of habitat extent can be achieved through inter-calibration when combined with a well designed environmental stratification [Jongman et al. 2006] and a habitat classification system such as the BioHab General Habitat Categories (GHC) system which is based on 'EO friendly' physiognomic characteristics. EBONE will investigate the success of inter-calibration applied on existing EO land cover maps which provide full coverage but also look at the inter-calibration of EO habitat maps of sample sites produced to increase the in situ samples in space and/or time. The advantage of the second approach is that it could allow for the introduction of strata specific EO mapping methods. In this context EBONE will look at the role of LIDAR and EO time-series analysis.

Species distribution models that incorporate in situ and EO derived information is another form of integration that could potentially deliver improved measures of the SEBI indicator 'Trends in abundance and distribution of selected species' [Gillespie et al. 2008]. Here both thematic (land cover maps calibrated to GHC observations) and quantitative EO derived information such as fragmentation and phenology metrics will be considered. One caveat is that distribution models highlight areas with high probabilities of a specific species occurring which does not necessarily mean the species in question will be found in all of the areas identified.

5. EBONE CONTRIBUTION TO EUROPEAN AND WORLD-WIDE PROGRAMS

5.1. GEO and GEO-BON

The Group on Earth Observations or GEO (www.earthobservations.org) is a partnership of 76 member nations and more than 50 NGOs, working to benefit society by improving the coordination of existing Earth observation data sets and implementing new observations and related products. It is designing a Global Earth Observing System of Systems (GEOSS) as the mechanism to achieve these goals. Biodiversity is one of the nine Societal Benefit Areas set forth by GEO as foci for its work. Thus, a Biodiversity Observation Network (GEO BON) is one of the first systems GEO is proposing for the GEOSS.

By facilitating and linking efforts of countries, international organizations, and individuals, GEO BON will contribute to the collection, management, sharing, and analysis of data on the status and trends of the world's biodiversity. It will also identify gaps in existing observation systems and promote mechanisms to fill them. The role of EBONE in this context is to act as a pilot for Europe that can be used by comparable initiatives in other continents.

The scope of GEO BON includes the terrestrial, freshwater, coastal, and open ocean marine components of biodiversity. Its definition of biodiversity encompasses genetic, species and ecosystem levels. In addition to collecting time series of observations on the presence, abundance and condition of elements of biodiversity at all of these levels, it will conduct limited analyses, such as change detection, trend analyses, forward projections, range interpolations and model-based estimations of the supply of ecosystem services. It will act in support of more detailed assessments undertaken by biodiversity and ecosystem assessment bodies. EBONE focuses on the terrestrial environment and an additional initiative for the marine environment is very much wanted.

The main users of GEO BON will likely be national governments (especially in relation to their obligations under biodiversity-related conventions) and their natural resource and biodiversity conservation agencies at national and regional levels, international organisations and the biodiversity-relevant treaty bodies, non-governmental organisations (both national and international) in the fields of biodiversity protection and natural resources management, and environmental and scientific research organisations both in and out of academia.

The EBONE project is the European contribution to GEO BON. It is developing a system of biodiversity observation at regional, national and European levels as a contribution to European reporting on biodiversity as well as to the GEOSS tasks on biodiversity and ecosystems. EBONE assesses existing approaches on validity and applicability starting in Europe, expanding to regions in Africa and seeking cooperation with projects in other continents.

5.2. LIFE-Watch

LifeWatch is an integrated approach for developing an advanced infrastructure for biodiversity research using a wide range of techniques. It will be significant in tackling one of the major challenges facing modern society, bringing together many disciplines and investigative techniques, all at the cutting edge of research, and will be how much research is done in the future. It intends to make massive biodiversity data sets available, searchable through an user interfaces; tens of thousands of users exploring these data and joining forces in virtual user groups. It also intends to give access to scientists and policy makers comparing and supplementing these data with even more data obtained from weather stations, satellites, biological collections from all over Europe.

The contribution that EBONE can provide to LifeWatch is the development of a ready available system for observation and data storage; not only for species, but also for habitats and supporting earth observation.

5.3. Extension of EBONE approach to Mediterranean regions outside Europe

The EBONE approach for Europe will need to be compatible with approaches at the world-wide level. Through a pilot for global Mediterranean systems EBONE will adapt the system that will be developed for Europe for Mediterranean and desert systems in test areas in Israel and South Africa as representative countries for this biome. This allows linking European approaches to Mediterranean and desert environment elsewhere in the world and allows testing of the methodology. A fundamental feature of the common approach developed for the definition of habitats in EBONE is that it is based on plant life forms that are also widely used basis for describing word biomes.

Within the EBONE project, two partners focus on Mediterranean-desert gradients, one in Israel and one in South Africa. CSIR, the South African partner, works from an existing Biota-Africa regional network similar to the LTER system, in which research stations are positioned regularly along the western coast of South Africa and into Namibia (see www.biota-africa.de). In this system there is heavy emphasis on remote sensing and detailed in situ biodiversity studies; the habitat mapping used in EBONE bridges a gap in this system.

In contrast, the Israeli partner INPA, is focused on conservation management at the habitat level, and is particularly concerned in effective habitat mapping. INPA and its partners in the Israel LTER system also have several research stations along a rainfall gradient from Mediterranean to desert regions. The stations are linked by a common protocol for monitoring the response to thinning and grazing by different groups of organisms in a statistically valid experimental block design. Some stations also carry out additional studies.

Two of these stations are currently in use for testing EBONE methodology; one in desert at the Avdat LTER site in the Negev Highlands, and the other in Mediterranean forest and maquis at Ramat HaNadiv near Mt. Carmel (for details on these LTER sites see http://aristo4bgu.bgu.ac.il/maarag/Default.aspx). Orthophoto coverage is used in all the Israel LTER sites, and coverage is good for the country. Ramat HaNadiv is a well established research site, and has an extensive monitoring program for plants, vertebrates, invertebrates, and geochemistry in place. In contrast, Avdat is a recently established LTER site mainly concerned with collecting information on plants and invertebrates.

True deserts are outside the range of European environments, so the methods and descriptors of habitats have to be modified for use in Israel. During the past year, suitable adaptations have been developed, which are currently being tested in the field by survey of five test squares in Avdat and five in Ramat HaNadiv.

Two workshops and field exercises have been held in the past year to introduce and train Israeli scientists in the use of the General Habitat Categories used for mapping habitats in EBONE. There is strong interest among Israeli conservation organizations to develop and use such a standardized habitat mapping system to enables sharing of data. The pool of participants in this training has therefore been drawn from conservation, forestry, rangeland, and basic ecological research organizations. There is developing consensus among these organizations that, with appropriate adaptation, EBONE procedure would be suitable for implementation throughout Israel. A third workshop is planned for November 2009, in which the results of the current habitat mapping experiments will be reported, and the range of interest will expand to the EO and biodiversity monitoring methodology used in EBONE.

The development of a unified habitat mapping methodology system is well under way in Israel, with discussion and adaptation now taking place daily. EO methods are not currently seen to be problematic. The weak part of the current situation in Israel is the methodology related to biodiversity. It is necessary to determine which measures of biodiversity are most meaningful and how these may be linked to habitat and EO data. The SEBI 2010 program has been introduced to the INPA and to the Israel LTER for consideration as a framework for biodiversity monitoring and is currently under discussion. As EBONE also works with the SEBI2010 biodiversity indicators, there will be much in common between the work programmes, designed to eventually lead to a standard system for the Mediterranean biome.

6. CONCLUSIONS

The main outcome of the project will be an integrated monitoring system based on key biodiversity indicators designed to be implemented within an institutional framework operating at the European level. This framework will provide continued access to indicator data for CBD reporting against the 2010 target and form the basis for the continued development of a European Biodiversity Observation system.

EBONE is a global pilot for international cooperation in biodiversity monitoring tackling the technical problems of harmonising approaches that differ in many ways:

- Topic: species, habitats and earth observation;
- Scale: from insects to migrating birds;
- Biogeography: linking Boreal, Mediterranean and Desert habitats and species;
- Organisation: trying to convince over 100 European agencies and an unknown number of NGOs to harmonise approaches.

This challenge has to be met for global reporting on biodiversity.

ACKNOWLEDGEMENTS

The EBONE project is a EU 7th Framework Programme, Theme 6 Environment, collaborative project No. 21322 in support of GEO.

REFERENCES

Bartholome, E., and A. Belward, GLC2000; a new approach to global land cover mapping from Earth Observation data, *International Journal of Remote Sensing*, 26, 1959-1977, 2005.

Bell, S., M. Marzano, J. Cent, H. Kobierska, D. Podjed, D. Vandzinskaite, H. Reinert, A. Armaitiene, M. Grodzińska-Jurczak, and R. Muršič, What counts? Volunteers and their organisations in the recording and monitoring of biodiversity. *Biodiversity Conservation,* 17, 3443-3454, 2008.

BirdLife International, Birds in the European Union: a status assessment. Wageningen, The Netherlands, BirdLife International, 2004.

Bock, M., P. Xofis, J. Mitchley, G. Rossner, and M. Wissen, Object-oriented methods for habitat mapping at multiple scales - Case studies from Northern Germany and Wye Downs, UK, *Journal for Nature Conservation,* 13, 75-89, 2005.

Delbart, N., T. Le Toan, L. Kergoat, and V. Fedotova, Remote sensing of spring phenology in boreal regions: A free of snow-effect method using NOAA-AVHRR and SPOT-VGT data (1982-2004), *Remote Sensing of Environment*, 101, 52-62, 2006.

Esseen, P.-A., A. Glimskär, G. Ståhl, and S. Sundquist, Field sampling protocol for National Inventory of Landscapes in Sweden (in Swedish). Dept. of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå, 222 pp., 2006.

Estreguil, C., P. Vogt, and K. Ostapowicz, European level assessment of the status and trends of forest spatial patterns*, Environmental Monitoring and Assessment Journal,* submitted.

European Environment Agency, Halting the loss of biodiversity by 2010: proposal for a first set of indicators to monitor progress in Europe. Office for Official Publications of the European Communities, 2007. http://reports.eea.europa.eu.

EuMon, EU-wide monitoring methods and systems of surveillance for species and habitats of community interest, 2009. http://eumon.ckff.si/

Friedl, M.A., D.K. McIver, J.C.F. Hodges, X.Y. Zhang, D. Muchoney, A.H. Strahler, C.E. Woodcock, S. Gopal, A. Schneider, A. Cooper, A. Baccini, F. Gao, and C. Schaaf, Global land cover mapping from MODIS: algorithms and early results, *Remote Sensing of Environment*, 83, 287-302, 2002.

Gregory, R.D., A. van Strien, P. Vorisek, A.W. Gmelig Meyling, D.G. Noble, R.P.B. Foppen, and D.W. Gibbons, Developing indicators for European birds, *Philos Trans R Soc Lond Ser B Biol Sci,* 360, 269–288, 2005.

Gillespie, T., G. Foody, D. Rocchini, A. Giorgi, and S. Saatchi, Measuring and modelling biodiversity from space, *Progress in Physical Geography*, 32, 203-221, 2008.

Haines-Young, R.H., C.J. Barr, H.I.J. Black, D.J. Briggs, R.G.H. Bunce, R.T. Clarke, A. Cooper, F.H. Dawson, L.G. Firbank, R.M. Fuller, M.T. Furse, M.K. Gillespie, R. Hill, M. Hornung, D.C. Howard, T. McCann, M.D. Morecroft, S. Petit, A.R.J. Sier, S.M. Smart, G.M. Smith, A.P. Stott, R.C. Stuart, and J.W. Watkins, Accounting for nature: assessing habitats in the UK countryside, DETR, London, 2000.

Henry, P.-Y., S. Lengyel, P. Nowicki, R. Julliard, J. Clobert, T. Čelik, B. Gruber, D.S. Schmeller, V. Babij, and K. Henle, Integrating ongoing biodiversity monitoring: potential benefits and methods, *Biodiversity Conservation,* 17, 3357-3382, 2008.

Heumann B, Seaquist J, Eklundh L and Jonsson P (2007). AVHRR derived phenological change in the Sahel and Soudan, Africa, 1982–2005. *Remote Sensing of Environment*, 108, 385 – 392

Hovestadt, T., and P. Nowicki, Process and measurement errors of population size: Their mutual effects on precision and bias of estimates for demographic parameters, *Biodiversity Conservation,* 17, 3417-3429, 2008.

Hill, R., A. Wilson, and M. George, Mapping tree species in temperate deciduous woodland using time-series multi-spectral data, *Journal of Applied Vegetation Science,* submitted.

Hill, R., and A. Thomson, Mapping woodland species composition and structure using airborne spectral and LiDAR data*, International Journal of Remote Sensing*, 26, 3763–3777, 2005.

Jakubauskas, M., D. Legates, and J. Kastens, Crop identification using harmonic analysis of time-series AVHRR NDVI data, *Computers and Electronics in Agriculture*, 37, 127-139, 2002.

Jongman, R.H.G., R.H.G. Bunce, M.J. Metzger, C.A. Mucher, and D.C. Howard, A statistical Environmental Stratification of Europe: objectives and applications. *Landscape Ecology,* 21:409-419, 2006.

Jurczak, M., and R. Muršič, What counts? Volunteers and their organisations in the recording and monitoring of biodiversity, *Biodiversity Conservation,* 17, 3443-3454, 2008.

Kull, T., M. Sammul, K. Kull, K. Lanno, K. Tali., B. Gruber., D. Schmeller, K. Henle, Necessity and reality of monitoring threatened European vascular plants, *Biodiversity Conservation,* 17, 3383-3402, 2008.

Lengyel, S., E. Déri, Z. Varga, R. Horváth, B. Tóthmérész, P.-Y. Henry, A. Kobler, L. Kutnar, V. Babij, A. Seliškar, C. Christia, E. Papastergiadou, B. Gruber, and K. Henle, Habitat monitoring in Europe: a description of current practices, *Biodiversity Conservation* 17, 3327-3339, 2008.

Lengyel, S., A. Kobler, L. Kutnar, E. Framstad, P.-Y. Henry, V. Babij, B. Gruber, D. Schmeller, and K. Henle, A review and a framework for the integration of biodiversity monitoring at the habitat level, *Biodiversity Conservation,* 17, 3341-3356, 2008.

Lillesand, T., and R. Kiefer, Remote Sensing and Image Interpretation, Third Edition, John Wiley and Sons Inc., 1994.

Loveland, T.R., and A.S. Belward, The IGBP-DIS global 1km land cover data set, DISCover: first results, *International Journal of Remote Sensing*, 18, 3289-3295, 1997.

Moody, A., and D. Johnson, Land-surface phenologies from AVHRR using the discrete fourier transform, *Remote Sensing of Environment*, 75, 305-323, 2001.

Oindo, B., and A. Skidmore, Interannual variability of NDVI and species richness in Kenya*, International Journal of Remote Sensing*, 23, 205-298, 2002.

Ortega, M., R.G.H. Bunce, J.M. García del Barrio, and R. Elena-Rosselló, The relative dependence of Spanish landscape pattern on environmental and geographical variables over time. Investigación Agraria: Sistemas y Recursos Forestales, 2008.

Paradella, W., M. Da Silva, N. Rosa, and C. Kushigbor, A geobotanical approach to the tropical rain forest environment of the Carajas mineral province (Amazon region, Brazil), based on digital TM-Landsat and DEM data, *International Journal of Remote Sensing*, 15, 1633-1648, 1994.

Schmeller, D.S.,European species and habitat monitoring: where are we now? *Biodiversity Conservation,* 17, 3321-3326, 2008.

Schmeller, D.S., P.-Y. Henry, R. Julliard, B. Gruber, J. Clobert, F. Dziock, S. Lengyel, P. Nowicki, E. Deri, E. Budrys, T. Kull, K. Tali, B. Bauch, J. Settele, C. Van Sway, A. Kobler, V. Babij, E. Papastergiadou, and K. Henle, Advantages of volunteer-based biodiversity monitoring in Europe, *Conservation Biology*, 1523-1739, 2008.

Thomson, A., R. Fuller, M. Yates, S. Brown, R. Cox, and R. Wadsworth, The use of airborne remote sensing for extensive mapping of intertidal sediments and saltmarshes in eastern England, *International Journal of Remote Sensing*, 24, 2717–2737, 2003.

Van Swaay, C.A.M., P. Nowicki, J. Settele, amd A. Strien, Butterfly Monitoring in Europe - methods, applications and perspectives, *Biodiversity Conservation,* 17, 3455-3469, 2008.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

EarthLookCZ - GMES data publication, combination and sharing on the web

P. Horak[a] **and M. Vlk**[b]

[a] *WIRELESSINFO, Cholínská 19, Litovel, Czech Republic (horak@wirelessinfo.cz)*

[b] *Help forest Ltd., Slovanská 21, Šumperk, Czech Republic (mavlk@helpforest.cz)*

Abstract: How to work with GMES data? The EarthLookCZ project looks for the ways, how to connect GMES data, how to publish own data and combine it with the others, how to create new services from available data and share them on the web … The technical output from the project, a prototype of the EarthLookCZ portal, is of course only one from the several possible answers to these questions. The solution of the project is based on the idea to implement several independent tools for geo-spatial data searching, data management and visualization into portal and to give a possibility to users to work with their own data together with data from external sources. Also an implementation of INSPIRE principles is one from the most important steps in portal development. EarthLookCZ portal presents new way of the GMES data utilization.

Keywords: GMES; INSPIRE; Spatial data; OGC; Core services.

1. INTRODUCTION

Project EarthLookCZ (www.earthlook.cz) is one of pilot projects developed in the Czech Republic in the ERA-STAR regions framework. The project undertaken by WIRELESSINFO association aims to verify the validity of spatial data infrastructure quality for GMES in the Czech Republic and is supported by the Ministry of Education, Youth and Sports. WIRELESSINFO association (www.wirelessinfo.cz) is a non-profitable consortium. We are a virtual research centre, which works on information technology development mainly in the areas of forestry, agriculture, environment and e-governance sectors.

One of the main goals of the EarthLookCZ project is to support the implementation of GMES in the Czech Republic. The report "Analysis of GMES Situation in the Czech Republic" was elaborated in 2007 as the first project output and serves as a background for the proposal of a "Prototype of GMES national portal" – the main technical output of the project. The analysis describes GMES activities at both European and national levels, GMES and INSPIRE interaction and GMES geo-data layers currently available in the Czech Republic governmental sector.

The proposal a new technological infrastructure is based on the implementation of ISO and OGC standards for data sharing and exchange. The objective of the solution is to provide a distribution system that will ensure access to distributed data and metadata through the GMES services. This solution, based on an earlier prototype (Early prototyping), will serve to verify the principles of GMES services catalogue at a national level. These services will be used in an international context with the EarthLookCZ project. The Catalogue portal GMES is one of the independent components of a comprehensive EarthlookCZ solution as a system for sharing and publishing of raster data. The GMES Data Portal provides data searching resources on the basis of their metadata records through structured requests. In the future, the portal will also include an editing function for creating metadata records. The Metadata and catalogue system will correspond to ISO standards 1919115/19119/19139 and will enable cascading searching facilities of other lists of standardized systems, as, for example, the GeoNetwork does.

2. GMES

Global Monitoring for Environment and Security is a European initiative aimed at providing stable, reliable and timely services in the fields of environment and security. This initiative is built on the basis of cooperation between the European Space Agency (ESA) and the European Commission. ESA ensures the implementation of various spacecraft components and coordination centres in Europe and the European Commission creates a basic concept, defines the priorities and manages the survey and development of services based on local data and data from remote sensing.

The GMES together with the planned Galileo satellite navigation system are the main pillars of the European space policy. The introduction of a comprehensive system that will include technology for remote sensing and local data collection and simple distribution methods, will hopefully contribute to the introduction of highly efficient services in various areas. These services exist today, but only at national or regional or field level without important joint coordination. From a pan-European point of view, the crucial thing is to ensure the possibility of using these services repeatedly, organizing and making them available as easily as possible. The first step has been done through the launch of the first "Core Services". It is very important to now find the best way in which we can move forward. For this reason, great emphasis is being placed upon the use of geospatial information and the harmonization and efficient management of spatial data across Europe. Also stress is being placed on the implementation of INSPIRE (Infrastructure for Spatial Information in Europe) as a guideline for infrastructure and management for the use and sharing of spatial data.

The GMES Structure

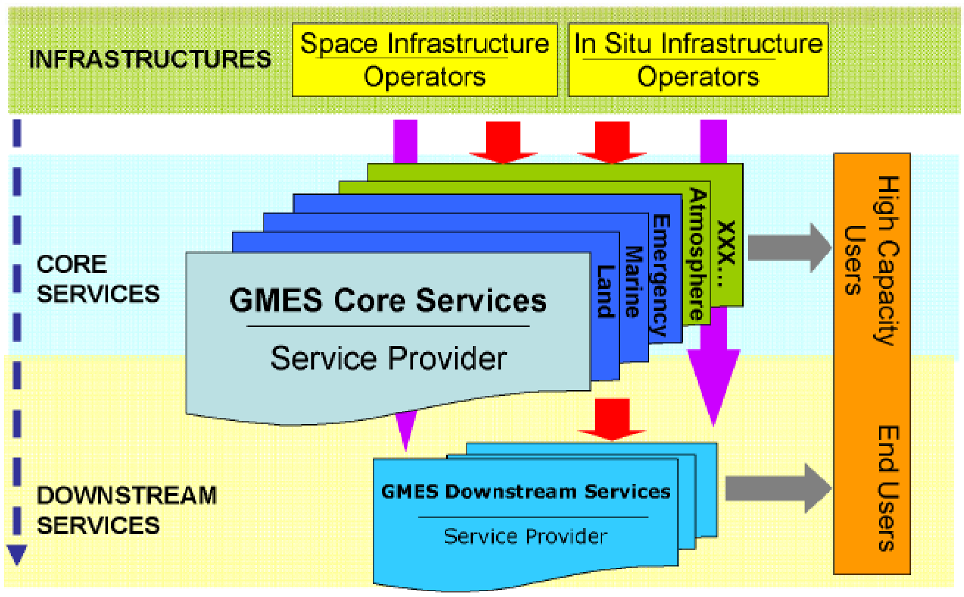


Figure 1. GMES structure

Core Services
- Provide standardized common multi-purpose information capacity for Europe;
- Link with European information needs

Downstream services
- Tailored for specific applications at local, regional, national levels (public or private use)
- EU should encourage and support the implementation of these service layers

As examples of the "Core Services" we should mention the Core Service Generic Landcover - a service developed within the framework of the integrated GMES project Geoland which provides thematically and geometrically corrected data land cover following the CORINE Land Cover. Next is the Service Core Service Bio-Geophysical Parameters, which provides basic information on the bio and geophysical attributes of the Earth's surface, such as the geophysical parameters of the vegetation activity or the water cycle.

3. INSPIRE

The Directive of the European Commission and Council “INSPIRE” - Infrastructure for Spatial Information in Europe- aims to create the legislative framework needed to build a European infrastructure for spatial information. Spatial data utilization is currently limited by the restrictive data policies of individual countries or involved subjects, as well as a lack of coordination. The biggest obstacles are the incomplete technical standards and their implementation, incompatible data and information systems, data fragments and the lack of data or data duplication in different places.

The basic principles of INSPIRE:

- Data should be collected once and maintained at a level where this can be provided most effectively
- It should be possible to seamlessly combine spatial data from different sources and share it between many users and applications
- Spatial data should be collected at one level of government and shared between all levels.
- Spatial data should be available on the condition that there are no restrictions on its usage.
- It should be easy to discover which spatial data is available, to evaluate its suitability and to know which conditions apply for its use

Activities in the areas of INSPIRE and GMES are among the main priorities of the European Commission in the field of information technology. GMES provides the acquisition of new data and its availability, INSPIRE defines the general requirements for the structure of spatial data and rules that allow sharing. Data interoperability, information sharing, data and metadata availability and the creation of new services present the main synergy requirements for GMES and INSPIRE.
The principles mentioned above are the backbone of the EarthLookCZ project solution.

4. THE PROTOTYPE OF THE NATIONAL GMES PORTAL

Currently, GMES data is available to users on the internet mainly as simple visualizations in the web map client. This solution makes data visualization possible, but the users have no possibility to use this data for their own purposes. A possible solution is to buy the “hard” data or use the data through standardized web services. Unfortunately only a small amount of GMES data is accessible through standardized web services now, but it seems that data sharing through web services will be the most important part of data exchange in the future.
The proposed solution of the GMES national portal in the EarthLookCZ project presents a way forward from the common methods for accessing data. The prototype of the portal is aimed at sharing and publishing the raster data (e.g. satellite imagery, orthophoto, geophysical measurements, climate data), vector data, thematic data on the environment, etc. The main innovation of the project is the ability of users to publish their own data sources on the Web and integrate their own data together with external data into new compositions. These map compositions may subsequently be published in the form of web services (Figure 2). Users also can integrate these new compositions into their own SW solutions. This concept of the GMES portal allows not only the common data visualization, but also new composition preparations by GIS specialists and data accessibility to other users, all within the portal solution.

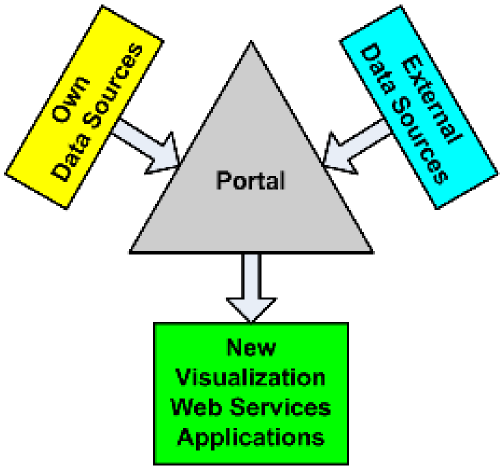


Figure 2. GMES portal concept

National GMES Portal enables an integration of monitoring data, that has been acquired using the space infrastructure (satellite imagery), and the In Situ infrastructure (ground measurements, aerial photographs) along with other data layers, such as general reference data (topographic maps) and thematic data (Landscape planning data). Also sources of GMES Core Services built within the basic structure of the European GMES will be accessible from the portal (Figure 3). These data sources will be accessible via web services.

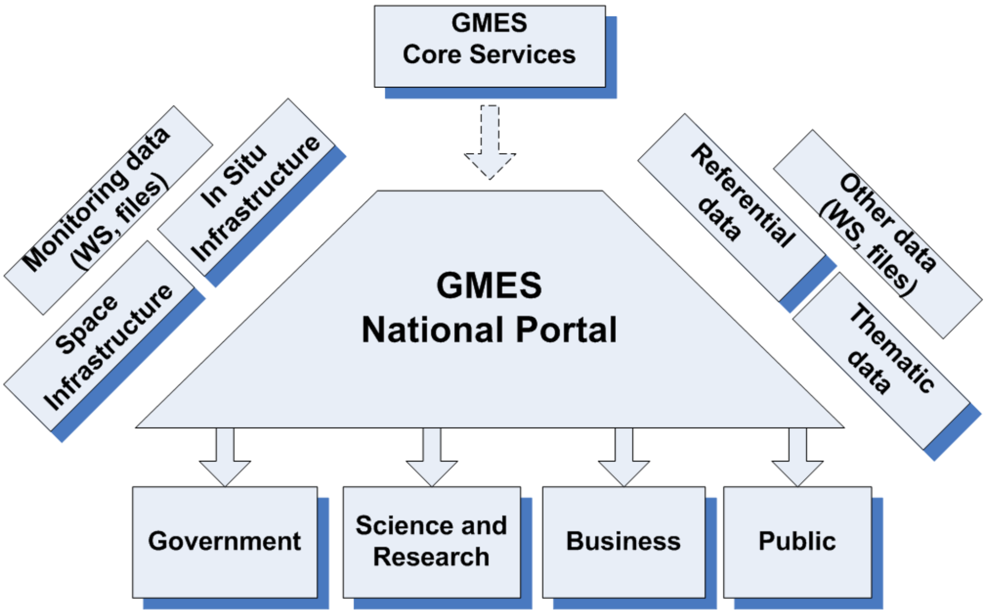


Figure 3. GMES portal functionality

Using the portal functionality, users are also able to upload their own data onto the server and publish it directly from the GMES portal. Users can combine different data sources, create new map compositions and make them available to other users. Availability "data about data" and map compositions is ensured through metadata and catalogue systems.

4.1 Overview of the basic functionality of the national GMES portal prototype:

- Data sources searching in the Metadata Catalogue
  - Searching of external data sources or own data on the server is provided by (?) metadata catalogue
- External GMES data accessing through web services
  - It is possible to use external GMES data if it is accessible through web service
  - Implementation of national and international data sources including GMES Core Services
- Own geo-data import onto portal and its publication on the web
  - User is able to import his own geo-data onto the server and offer this data to others as a standardized web service
  - Import of files into file repository or geo-database
- Management of the both internal and external data
  - It is possible to combine external and internal data and create new map compositions from them
- Data visualization in web map client
  - All internal data, external data and new compositions are visible in a map application built into the portal
- Creating new web services
  - Internal data and new map compositions can be published as OGC web services
- Component generating for outside web pages
  - The system makes it possible to generate a map application with an open API interface; user can integrate this with other applications or web sites
- Metadata generation
  - Internal data and new compositions can be equipped with standardized metadata records Photographs should only be used if essential to the clarity of the paper. If used they must be black and white with clear contrast and highly glossed.

4.2 GMES Core Services and portal interactions

Currently, Core Services outputs are seldom available in web service formats, therefore the implementation of these services into the portal is problematic. Nevertheless, Core Services data is supposed to be more easily accessible as WMS, WFS or other OGC services in the future; hereby the data could be joined to new compositions on the EarthLookCZ portal. The first implementation is hoped to be included in the next version of the portal this year.

4.3 The technological concept of the portal prototype

The concept of the portal is divided into 3 main parts:

- Sample map applications and metadata catalogue – public section of the portal where users can work with pre-defined GMES map applications or search for GMES data sources in metadata catalogue.
- GMES data management – the most innovative part of the portal; registered users are able to import their own data onto the portal, create new map compositions, integrate external GMES sources through web services and make them available to other users as a new web service
- Sources of information about GMES – links to the most important GMES web sites and public GMES web sources

Currently the first version of the portal prototype is available with very provisional functionality and amount of sources (Figure 4). The final version of the prototype is expected at the end of 2009.

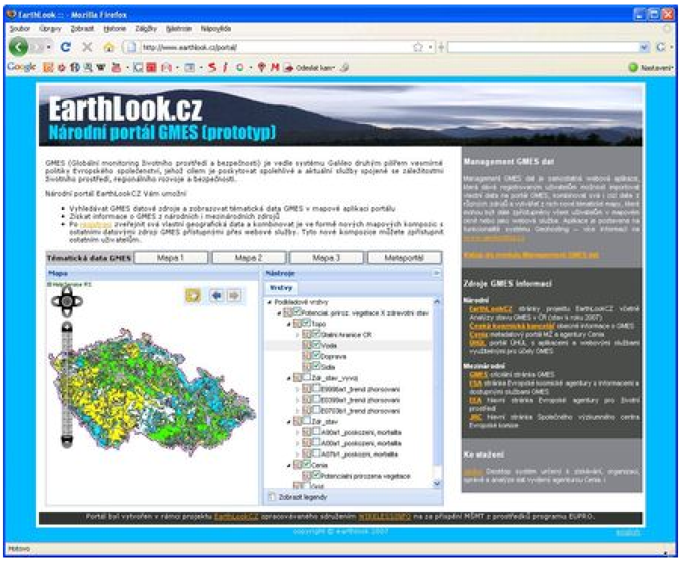


Figure 4. Prototype of the GMES portal

The solution of GMES Data Management is based on the GeoHosting system principles. GeoHosting was developed by WIRELESSINFO members and offers services supporting the creation of a spatial data sharing system with the possibility of publishing data for any user who has access to the Web. The system is based upon open formats and is open for interaction with other SDI platforms. This model supports the re-usability of components and easy building of new applications or their modification. The goal of the design is to re-use existing tools and to define interoperable interfaces for these tools, which will therefore create the possibility of re-using these tools as a part of other applications. For the selection of the components it is important to be familiar with Open Source solutions, but only on such systems which support Open Standards (OGC and ISO) (Figure 5).

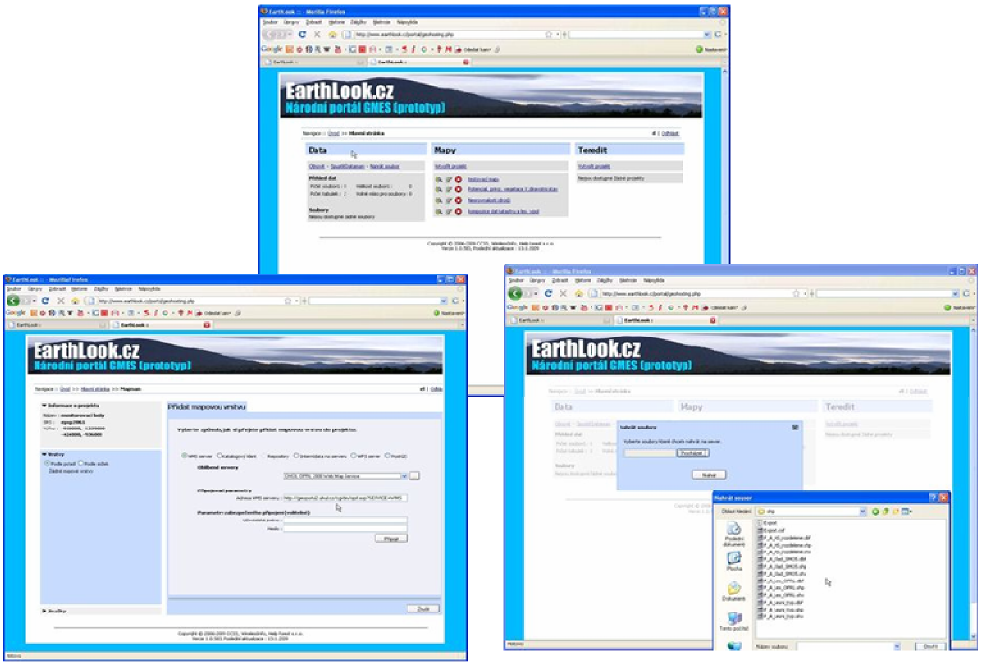


Figure 5. GMES data management

Currently, Data Import and Map Management components are implemented into GMES portal prototype(s).
Data Import is an application for the management of spatial data. It supports the management of data in databases or files. It supports the export and import of this data and also publishing and updating of related metadata. In databases, it is possible to store both, vector and raster data, including their attributes. Also for file oriented storage, it supports both vector and raster data. From raster formats, it currently supports IFF/GeoTIFF, JPEG,GIF, PNG, BMP, ECW, from vector formats ESRI Shapefile, DGN, DWG, GML

The Map Management application is a software tool for users who want to publish or create new map projects and compositions. It supports the publication of spatial composition from locally stored data (fields or databases-stored in the Data Import application), with external WMS, WFS data services. It supports visualization in web browsers using such clients as OpenLayers, GoogleMaps, DHTML client, Desktop viewer, GoogleEarth, DIS Janitor. It can also publish data as OGC WebMapService (WMS), OGC WebFeatureService (WFS). All published data is also connected with metadata stored in Micka metadata catalogue.

5. CONCLUSION

The prototype of the EarthLookCZ portal illustrates possible ways for GMES data utilization in accordance with INSPIRE principles. Users of the portal can search for information regarding GMES on both a national and international level. When they are registered, they are able to actively work with available external data sources, but they can also use their own data for integration with external sources, create new map compositions and make them available to other users. Publication of their own data or map compositions is joined with synchronous publication of their metadata; this functionality ensures, that the new data is searchable for other users. An interaction with GMES core Services is desirable; practical examples should be available in the last period of the EarthLookCZ project.

REFERENCES

Annoni, A., INSPIRE / GMES / GEOSS, presentation.
COMMUNICATION FROM THE COMMISSION TO THE COUNCIL AND THE EUROPEAN PARLIAMENT, Global Monitoring for Environment and Security (GMES): From Concept to Reality {SEC(2005)}.
Global Monitoring for Environment and Security, Final Report for the GMES Initial Period (2001-2003).
EarthLookCZ, report D12 - Specifikace GMES serveru, 2008.
EarthLookCZ, report D18 - Prototyp1, základní funkcionalita, 2008.
Harris, R., Professor - GMES AND INSPIRE - Legal and Policy Issues
http://www.cenia.cz.
http://www.czechspace.cz, http://www.czechspace.cz/cs/gmes.
http://www.earthlook.cz/portal.
http://www.gmes.info.
http://www.opengeospatial.org.
SDĚLENÍ KOMISE RADĚ A EVROPSKÉMU PARLAMENTU, Evropská vesmírná politika, {SEK(2007) 504}, {SEK(2007) 505}, {SEK(2007) 506}.
SMĚRNICE EVROPSKÉHO PARLAMENTU A RADY o zřízení Infrastruktury pro prostorové informace v Evropském společenství (INSPIRE) PE-CONS 3685/2006.
Šobra, J., GMES, presentation.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

DYVINE – Integrating a forest of cameras in urban environment

P. CHROBOCINSKI – G.VARANGOT
EADS Defence and Security – France
Philippe.chrobocinski@eads.com – guilhem.varangot@eads.com

Abstract: In this day and age where hundreds of cameras are being installed in every city, the main problem that arises is how to make the best of the huge number of streams of images now available, in order to enhance the security and safety of the citizens. The existing systems are based on closed-circuit television (CCTV) principle where the images of each camera are displayed on individual screens, which necessitate overgrowing number of operators who have to be constantly vigilant. Moreover, these CCTVs are operated by numerous independent organisations and the monitoring of the cities is therefore highly fragmented.

The concept of DYVINE is to design a federative system which is able at any time to integrate all the cameras of the town, whatever their belonging, complete them with other cameras, fixed or mobile, in situ or airborne, to provide the risk management agencies with an accurate and up-to-date situation picture of the relevant events. This result has been be achieved by developing innovative and robust automatic video content analysis modules which will highly lighten the tasks of the operators and by providing wireless communication solutions that will provide increase resilience to any type of aggression.

***Keywords**:* Cameras; video-processing; integration; wireless communications.

1. BODY OF PAPER

1.1 Introduction

How to fully harness in times of crisis the combined power of the thousands of cameras and sensors that populate the land? That is the question that DYVINE (Dynamic Visual Networks), a project funded by the EC’s Sixth Framework Programme, has been endeavouring to answer since September 2006. The solution could be highly attractive for contingency responders all around the world.

DYVINE acknowledges that visual sensors such as cameras are both affordable and widely used. However, at times of crisis – when access to these sensors would actually be useful – they are often not accessible by the people who need them the most because they are stand alone systems often owned by private companies or separate agencies.

1.2 The development

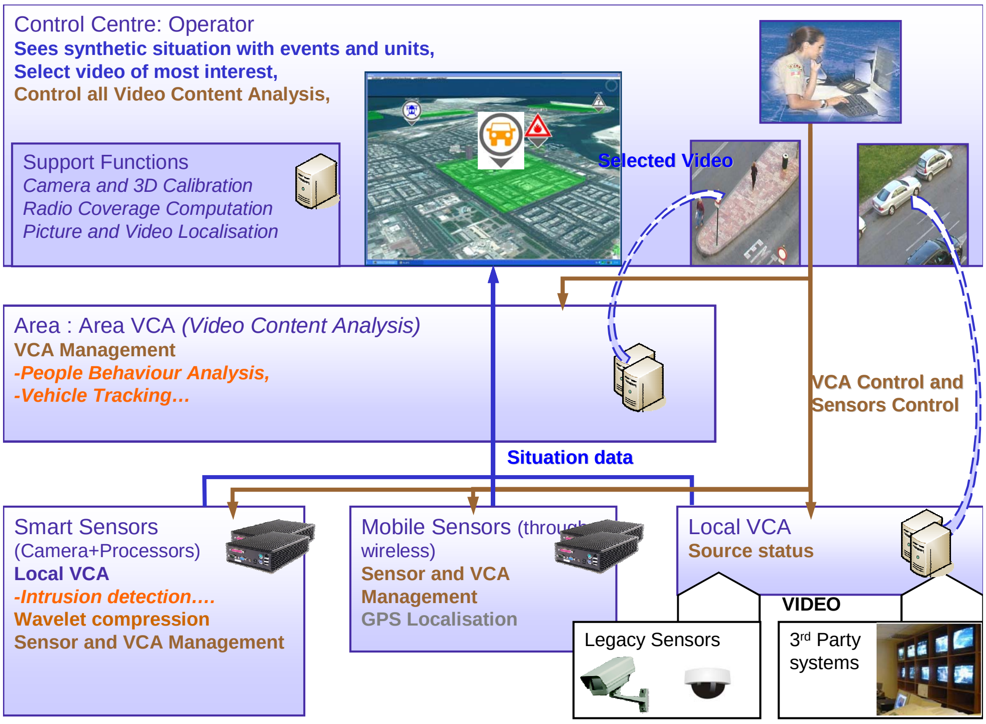
DYVINE is a flexible solution that can be used both during major disasters and day-to-day emergencies and citizen security.

To the consortium's knowledge, DYVINE and the sheer number of amalgamation capabilities being researched are unprecedented, which is why 12 partners were necessary to form the consortium.

Among the partners; EADS is designing the IT architecture; Martec (France) is developing the interface between the sensors and the system; Realviz (France) specialises in image processing; Surrey University is working on WiMAX integration; Valencia University on body worn cameras integration; and EPFL (Switzerland) on 3D processing for locating objectives, amongst others.

There have been three main challenges. First of all the building of an architecture capable of integrating into one system a large number of cameras – there could as many as four thousand in a city. Secondly, the ability to integrate new cameras such as body worn cameras, cameras on cars or new sensors into the system during times of crisis. The third challenge has been to optimise the video processing to allow real time processing.

The architecture for the integration of the sensors is presented in the following picture.



1.3 The trials

A successful first trial has taken place in Segrate, Italy, in November 2007, demonstrating common operational picture, video stitching, and video compression in an urban scenario with an important truck accident generating huge troubles in the city.

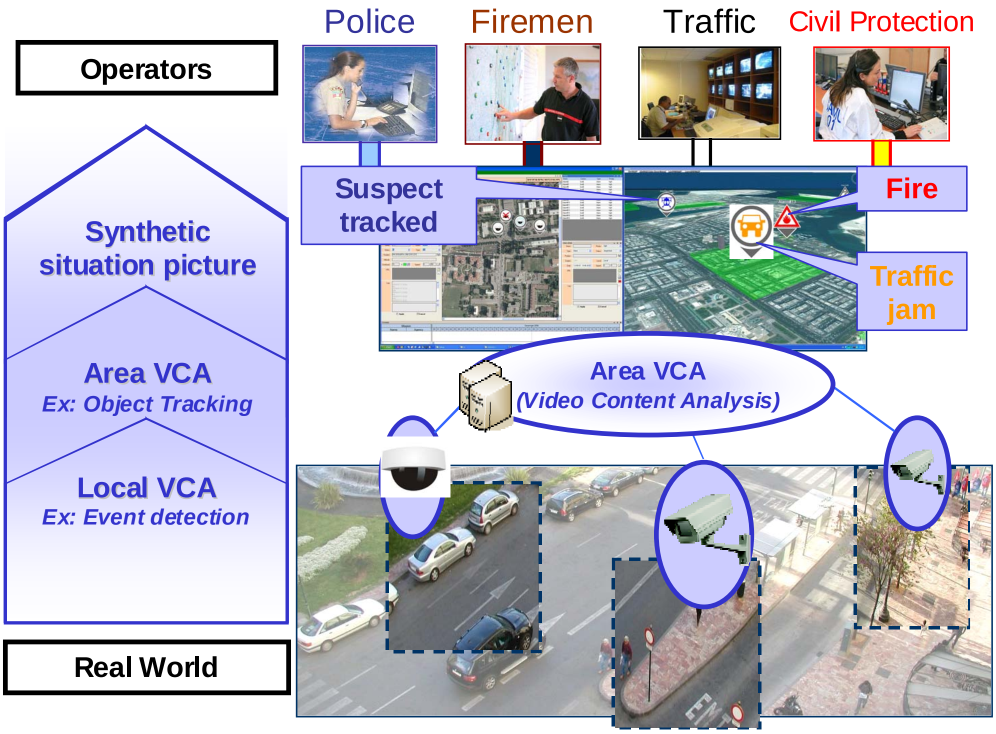
In the first step, a traffic jam was detected automatically via live video analysis and data fusion. The truck accident event was detected and shown to the operator; a subsequent fire on the truck was then also detected and reported; patrol car attendance shown on the system; and live images transmitted by the first responder were also shown on the system. The fusion and correlation of all the alarms were demonstrated on the overall picture.

In the third step, a hazardous area was outlined, demonstrating how the system could define a security perimeter, detect entry points; and observe moving people and objects in the perimeter through a tracking system.

The trial proved that DYVINE can be deployed easily and can interface with legacy surveillance systems without major impact on its architecture. The demonstration in Italy reached all the objectives with a realistic scenario – mostly with real events and data – in which the system operated seamlessly and with the expected results.
The invited operational end-users of the trials successfully validated the general concept of DYVINE, and provided valuable inputs for the next phases, in terms of adaptability to the various agencies way of working and in terms of system operability.

A second and final trial took place in July 2008 in Valencia, Spain, the culmination of the two-year project.

The trial in Valencia demonstrated all the technical capabilities of DYVINE in a scenario covering all the possible configurations of the system. Starting with day-to-day operation by the city traffic management, it escalated after a disaster to a full configuration integrating all the sensor networks and additional mobile cameras to provide an accurate real-time picture to the responders (Police and Civil Protection).



1.4 The future

It is now time to take DYVINE to the open market, and the consortium believes that having been designed to be highly affordable, it will be attractive. Our aim was to develop an affordable model to install on top of existing systems. Obviously if there are no existing camera networks it would be expensive to install. If we are talking about installing smart layers inside a fully equipped city then it is not costly.

Various configurations are available depending on the national contingency structure. Valencia city hall, for example, is in charge of all civil protection agencies, so in this case the city hall would be the owner of the system. In the case of France it is more complex because the local police have very limited responsibility, so here perhaps it would be the national police.

Spain would be a good example of a suitable end user, because common emergency centres have been developed here for 112 emergency calls. In the typical Spanish 112 room you have the legacy systems of the police, ambulance and fire. You can think of DYVINE as the layer that combines all that information and synthesises it. In a real scenario, it could mean for example a fireman being able to access live video footage from the cameras of a commercial shopping centre. The idea for this type of customer is that they can deploy selected parts of DYVINE to enable connection to CCTVs.

All has not been smooth and some challenges still remain. Data protection is still a problematic issue. The consortium had to take into account both European and national laws. The use of systems that can help identifying a person are subject to strict legal conditions. And secondly, connecting different agencies together means that you should be able to control what is done with each one of the entering video footage, originally managed by these distinct agencies.

There is a need within the system some control to ensure that all data is exploited in the right way. This area was not in the DYVINE framework for development but an exhaustive study has been performed by KU Leuven to identify all the relevant issues that need to be addressed when implementing such a system. Some solutions are suggested, but the main headache is that the data protection legislations have been enacted 10 years ago and are thus not adapted to systems such as the one developed by DYVINE. Specific CCTV legislations adopted in some countries does not deal with the specific problems raised by such systems neither.

The 12-partner consortium of DYVINE aimed to support the risk management cycle via the integration of thousands of fixed/mobile/airborne video sensors, providing a resilient communications system that will support their use in real time.

The outcome of DYVINE will shortly be on offer to the world, either as a whole solution or in segments, depending on the legacy systems already in a city or area. The consortium will be pitching this to many different parties: government – mostly for border control and security systems, City halls for Urban Security; but we want to talk to people responsible for critical infrastructure, airports, ports, nuclear plants.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

OneGeology-Europe: Making European geoscience spatial data accessible

Ian Jackson
British Geological Survey, NG12 5GG, UK (ij@bgs.ac.uk)

Abstract: Rich geological data assets exist in the Geological Surveys of each country of the world, including each European Member State, but they are difficult to discover and are not interoperable. For those outside the Survey, or nation they are not easy to obtain, to understand, or use. Geological data are essential to the prediction and mitigation of landslides, subsidence, earthquakes, flooding and pollution. Geology is a key dataset in INSPIRE. It is needed for the Groundwater and Soils Directives, GMES and GEOSS. OneGeology-Europe is making geological spatial data held by the Geological Surveys of Europe discoverable and accessible and will see Europe play a leading role in the global OneGeology initiative. The project is accelerating the development and deployment of a nascent international interchange standard for geology, GeoSciML, enabling the sharing of data within and beyond the geological community. OneGeology-Europe is facilitating re-use of geological data by a wide spectrum of public and private sector users. It is addressing the licensing and multilingual aspects of access and will help move geological knowledge closer to the end-user where it will have greater societal impact. The project will provide examples of best practice in the delivery of high resolution digital geological spatial data to users, e.g. in the insurance, property, engineering, mineral resource and environmental sectors. Through the project Europe is taking a leading role in the development of a global geoscience SDI and making progress towards INSPIRE goals. The project will deliver: an interoperable geology spatial dataset at 1:1 million for all the EU; higher resolution applied geological spatial data services for several Member States; a multilingual discovery portal; a robust OGC compliant data model, schema/mark-up language; a web portal providing multilingual access to the data; best practice examples of the delivery of geological data to a range of users; best practice licensing guidance; and last but not least exchange of science, informatics and business skills and experience across the EU and globally.

***Keywords**:* geology; geoscience; hazard; resources; interoperability

1. INTRODUCTION

Rich geological data assets exist in the Geological Survey of each individual Member State but they are difficult to discover, and are not interoperable. For those outside the Geological Survey, or nation they are not easy to obtain, to understand, or to use. Geological spatial data are, for example, essential to the prediction and mitigation of landslides, subsidence, earthquakes, flooding and pollution. These issues are global in nature and their profile has been raised by the United Nations International Year of Planet Earth 2008, OneGeology global initiative. OneGeology-Europe will accelerate the pace of this initiative within the EU and make a significant contribution to the global venture. Geology is a key dataset in INSPIRE (Annex II); it is also fundamental to the Annex III Themes of natural risk zones, energy and mineral resources. It is needed for the Groundwater and Soils protection Directives, the GMES and GEOSS programmes and for the development of the Shared Environmental Information System (SEIS), for Europe. OneGeology-Europe will make geological spatial data held by the geological surveys of Europe more easily discoverable and accessible via the Web.

2. OBJECTIVES

OneGeology-Europe is bringing together a web-accessible, interoperable geological spatial dataset for the whole of Europe at 1:1 million scale based on existing data held by the pan-European Geological Surveys. It is developing a harmonised specification for basic geological map data and will make significant progress towards harmonising the dataset (an essential first step to addressing harmonisation at higher data resolutions). It is accelerating the development and deployment of a nascent international interchange standard for geological data – GeoSciML, which will enable the sharing and exchange of the data within and beyond the geological community within Europe and globally. It will also facilitate re-use and addition of value by a wide spectrum of users in the public and private sector and identify, document and disseminate strategies for the reduction of technical and business barriers to re-use. It plans to address the multilingual aspects of access through a multilingual discovery portal. In identifying and raising awareness in the user and provider communities it will move geological knowledge closer to the end-user where it will have greater societal impact and ensure fuller exploitation of a key data resource gathered at huge public expense. The project will provide examples of best practice in the delivery of digital geological spatial data to users, e.g. in the insurance, property, engineering, planning, mineral resource and environmental sectors. Last but not least, OneGeology-Europe will see Europe playing a leading and pivotal role globally in the development of a geoscience SDI – a major and significant contribution to INSPIRE.

OneGeology-Europe is addressing head-on the challenges of interoperability and open standards. It will improve access to, and exploitation of, rich and relevant digital content. In bringing together and raising awareness amongst an extensive existing network – Geological Surveys who are pan-European providers with proven capacity and stability, plus users and stakeholders from a cross section of key sectors - it will assist in the dissemination of best practice, while at the same time identifying and addressing weaknesses, barriers, gaps and opportunities. While the problems and opportunities geology raises are trans-national many of the issues are currently being tackled on a local basis and in a disconnected way by individual nations. The project is creating synergies and aspires to facilitate the development of added value products and services across the European Community. A strong communication and awareness-raising component coupled with the enhancement of well established international networks is ensuring that the profile of the project is global and prominent. The long-term mandate and record of the individual geological surveys is a guarantee of the sustainability and future trajectory of the project outcomes.

3. GEOLOGICAL INFORMATION IN EUROPE

Finding solutions to the challenges of landslides and earthquakes, minerals and mining, water supply and flooding, pollution and erosion, and not least, climate change and energy supply, depend absolutely on spatial geological data. Like most things environmental, few of these challenges respect national frontiers and if we want to assess and address these environmental challenges at a European scale then we need access to European geological data too. Rich geological data does exist in each national geological survey, but when it is available (and in many instances in is exceptionally difficult to discover), then it exists in different formats and via different services, with different access conditions. There is a need to make these key datasets discoverable in a European, as opposed to just a national, dimension. Equally important is the need to be able to share and integrate data from different national providers – interoperability and harmonization is something that at the moment is not possible without expensive ad hoc data conditioning. Lastly, in several countries excellent skills and practices in data discovery and delivery and exploitation have been developed, often through local or national provider-user partnerships (e.g. ISO standard metadata discovery portals, real-time viewing and download services, client-driven services for the property and insurance sectors). In other countries, whilst the needs are virtually identical, the services and functions are much less well developed; the national and international user's needs are not met and the potential the provider has is unfulfilled. The complementary exemplary practices that exist need to be identified, shared and spread to improve the availability and re-use of geological data and provide conditions for development of added value public and commercial services across Europe.

Geological data that are currently available have a reasonable national user base, but even then in most Member States it tends to be strong in one or two sectors (mineral and energy resources and civil engineering) and with notable exceptions, commonly under-utilised in others (eg insurance, environmental assessment and protection, landscape and heritage) where it could add significant value across Europe. Thus digital geological data has considerable potential capacity for re-use by local, regional, national and European public authorities, and citizens and SMEs, throughout Europe, provided it is made discoverable, readily available, and interoperable and has clearer and more consistent use conditions. Virtually all geological data is spatially related and the project will contribute directly to the objectives of INSPIRE, where geology is a key Annex II dataset. Geology is also fundamental to the Annex III Themes of natural risk zones, energy and mineral resources. The project will assist in the development and deployment of the specifications for data, metadata, delivery services and data sharing. The Geological Survey data providers are all public sector bodies and thus the project is supporting Community policy in the Re-use of Public Sector Information Directive. The project will tangibly support the Environmental Information Directive and will enable a more complete and coherent implementation of Groundwater and Soils protection Directives, the upcoming Mine Waste Directive, in addition to GMES, GEOSS and SEIS.

OneGeology-Europe has 29 partners within the consortium from 20 nations. There are 20 Geological Surveys (data providers), 7 representatives of the user community, an expert on legal aspects of geographical data access and the umbrella organisation for all the Geological Surveys of Europe (EuroGeoSurveys). In addition there are organisations who are unable to be formal members of the consortium at this time but who have offered support; including several Geological Surveys who will provide data and 8 users and organisations who have agreed to be involved and provide guidance.

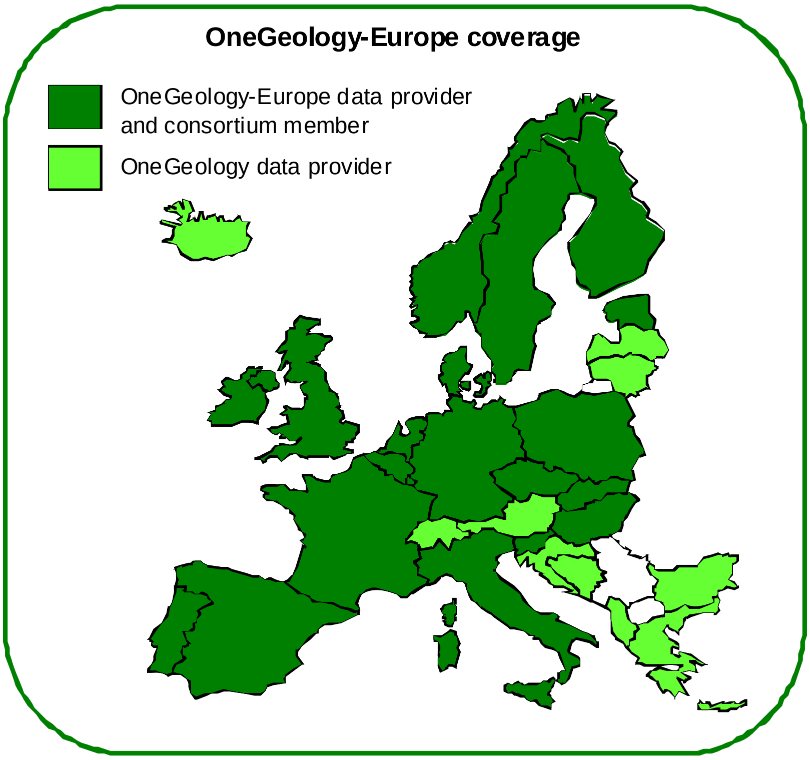


Figure 1. OneGeology-Europe coverage

4. USES AND USERS

An extremely wide range of organisations in commercial and public sector domains uses geological information. It is also used by many citizens through their leisure pursuits or education. Geological Survey data has been utilised by the resource and civil engineering sectors for 2 centuries in Europe and continues to be in demand today (e.g. minerals, roads and major structures). It is increasingly used by the planning and environmental sectors. The sustained existence of Geological Surveys in every nation of Europe is tangible evidence of the ongoing demand for this information. The fact that geology is a priority theme in INSPIRE (Annex II) is further verification. Specific evidence of demand includes: over 800000 environmental reports containing geological survey data generated for the property sector every year in the UK; the French Survey, BRGM, had 2.5 million visits to its InfoTerre, risk and pollution websites in six months; the Geological Survey of Denmark and Greenland's website received 1.8 million visits in 2006, predominantly related to access to hydrogeological (groundwater) data. In November 2006 a Pricewaterhouse Coopers report on the economic benefits of environmental science in the UK specifically concluded that subsidence data supplied by the British Geological Survey could reduce insurance industry costs of subsidence by between €4.3 million and €15.8 million per annum in the UK alone. These are specific examples, many others exist, but the fact is that because they remain disparate and national, Europe as a whole is not benefiting from the scale-up of such exemplars of data access and exploitation

In many individual nations, and across Europe as a whole users require consistent, quality assured, usable (understandable) geological data. Those users who operate trans-nationally would like, at a minimum, interoperable data and preferably harmonised data if possible. Consistency of access to the data in terms of clear, simple, transparent and fair licensing is also a pressing unmet demand. The data and portal to be developed and made available is intended to improve the accessibility of high quality information for all citizens of Europe – improving the equality of opportunity in terms of their access to diverse national information holdings. The information which will be accessible for a broad spectrum of purposes – education, hazard mitigation, mineral and energy resource exploration, groundwater protection, civil engineering, land and property development, planning and policy making, tourism and insurance. The project is likely to reveal previously "hidden" public data sources to governments, companies and citizens alike and in doing so aspires to decrease the cost of data sourcing and stimulate enhancement of the quality of services in the public and private sectors. Through transfer of know-how it will make essential but often esoteric, science data more understandable and useful. All these factors will lead to a better contribution by the data providers and government and commercial users to EC economic development and social objectives.

5 WORK PLAN AND RESULTS

OneGeology-Europe is a large and ambitious project, however, its size and extent and the internationally applicable goals have ensured it has significant profile and momentum. The project is broken down into 10 Work Packages (WP); these are:

WP1. Project management
WP2. User cases and best practice in meeting stakeholder needs
WP3. 1:1 million pan-European geological data specification, identification and sourcing
WP4. Data inventory and multilingual discovery metadata catalogue
WP5. Informatics specification, data model, interoperability and standards
WP6. Web portal and Registry development and implementation
WP7. Access and licensing protocols
WP8. Communication, dissemination and awareness
WP9. Access to high resolution geological & applied datasets in national & cross border situations
WP10. Liaison with related groups and initiatives

Work Package 3 merits amplification. It constitutes a major and essential piece of pan-European work, which needs a coherent approach to ensure integration. Work Package 3 is delivering a specification for geological spatial data and an interoperable 1:1 million scale dataset for the continent. Based on the specification it will identify the generic and specific geometric and semantic harmonisation issues. It will then coordinate the “mapping” of these existing national datasets to this specification to make significant progress towards a harmonised dataset – a critical step towards INSPIRE goals. The standards, architecture and framework developed can then be “upscaled” and progressively deployed for higher resolution geological data. The full list of deliverables for the project is listed in Table 1.

1	an interoperable geology spatial dataset at 1:1 million scale for all onshore EU (specifically composition and structure of the surface geology)
2	a scientific and informatics specification for the harmonisation of geological data at this resolution and progress towards a harmonised European dataset
3	a view service providing access to best practice high resolution geological spatial data services for 6 Member States (specifically composition and structure of the bedrock and superficial deposits, and thematic, added value, data on hazards and natural resources)
4	2-4 pilot and case studies on cross-border delivery of harmonised high resolution data access
5	multilingual discovery metadata for all data provider participants’ geological and applied map data
6	a robust data model, schema and mark-up language for the geosciences, which is OGC compliant and documented and deployed across the EU and internationally
7	a web portal providing easy multilingual access to the above data and examples of user-focused web services
8	documented best practice examples of the delivery of geological data to the range of users (available for download and as high quality published book)
9	guidance and proposed code of practice on licencing and clearing arrangements facilitating re-use of geological spatial data
10	exchange of science, technology, informatics and communication skills and experience across the EU and globally

Table 1. OneGeology-Europe deliverables

6. THE GLOBAL CONTEXT

On 6 August 2008, at the opening of the 33rd International Geological Congress in Oslo, Simon Winchester, author of the best-selling book about the English geologist, William Smith and his “Map that changed the world”, launched the web portal of a global project called OneGeology. Had you used the term “OneGeology” 32 months earlier – or typed it into Google - you would have registered a blank. The project and its name did not exist. OneGeology is now a major global venture - a contribution of the geological surveys of the world to the United Nations International Year of Planet Earth. This UN Year provided the stimulus to begin the creation of an interoperable digital geological dataset of the planet at 1:1 million scale. This is a multi-lateral and multi-national project that mobilises geological surveys to make an ongoing contribution and act as sustainable data providers of a global dataset. This vehicle of creating a tangible geological map is also being used to accelerate progress of an emerging global geoscience data model and interchange standard but perhaps most importantly the project is transferring know-how to developing countries and reducing the length and expense of their learning curve, while at the same time allowing them to serve maps and data that can attract interest and investment.

Following a kick-off meeting in Brighton, UK in March 2006, an international team from geological surveys around the world began work on these goals and in the next 18 months made astounding progress. They have made significant amounts of geological map data accessible – currently 96 nations are participating and up to 40 of these are serving data. They delivered a web map portal and the protocols, registries and technology to “harvest” and serve data from around the world. They exchanged know-how and produced guidance ("cookbooks") and provided support so that any geological survey can participate and serve their data. They moved forward and raised the profile of a crucial data model and interoperability standard – GeoSciML.

The technology used to achieve OneGeology's results to date is not complex, but it in terms of the scale of the deployment it is world leading. A basic principle of OneGeology (global) is that it must be inclusive and open to all geological surveys to participate, regardless of development status and the project has devised protocols and systems to ensure this. OneGeology is thus open to those who currently possess only traditional paper geological maps, and to those operating sophisticated web mapping systems. The end-user does not require specialist software, only access to the Internet via a web browser. In this first phase OneGeology is delivering digital geological map data from participating nations using Web Map Services (WMS). This is a distributed, dynamic and sustainable model, which leaves the data where it is best looked after and updated; that is with the provider nations. Each survey either registers its web service with the OneGeology portal or works with a partner survey (a "buddy") to serve that data. OneGeology technology is compliant with the international Open Geospatial Consortium (OGC) Web Map Service standard. Geological surveys may use a variety of software (e.g. MapServer) to serve their data. The Portal displays the map data served by each country and provides users with the ability to zoom, pan, switch map data on and off, change its opacity and even transfer it to Google Earth.

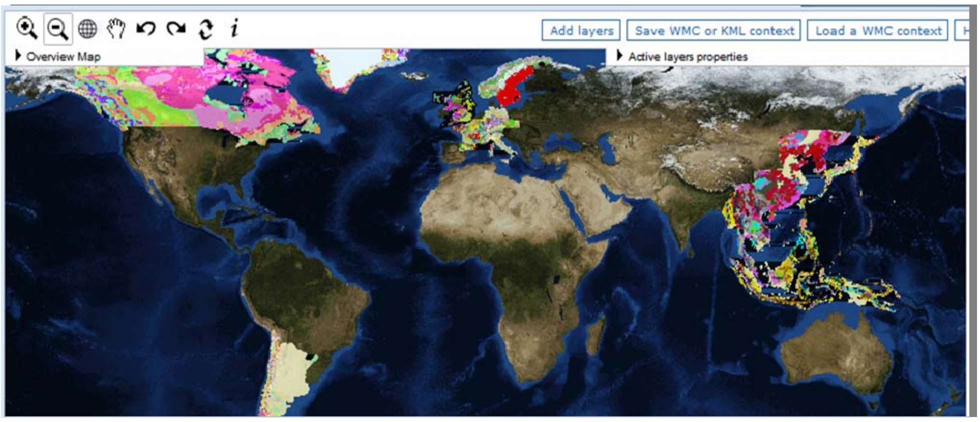


Figure 2. Extract from screenshot of the OneGeology portal

The outputs above are not the limit of the project achievements. In delivering its portal and technical protocols OneGeology has progressed something major global and regional bodies (including the United Nations and the European Union) have been advocating for some years – the creation of a spatial data infrastructure for planning and policy-making. The project has now been unanimously endorsed by the Directors of the geological surveys of the world meeting in Oslo and is providing a tangible catalyst for future collective and coherent action by surveys. Google references to OneGeology grew from 4000 on 1 August 2008 to over 220 000 on 15 August; it is not the size of this number alone that excites; when you look more closely at some of these web pages you see the way that "liberating" the data has allowed others to innovate and use their imaginations – from new teaching resources for geography students, to animated mash-ups and fly-throughs of Mount Fuji. The outreach has not stopped at the science and academic community; media interest in OneGeology has been extraordinary – over 700 articles and broadcasts worldwide in 4 weeks, from Nature to Vatican Radio, each in its own way presenting an opportunity to describe to audiences, who we would usually never reach, why geology is important to society - and never more important at a time of intense interest in energy needs and the impact of climate change.

The OneGeology initiative has made progress in funding too. In addition to the OneGeology-Europe project, in the USA the National Science Foundation is providing almost $700 000 for a similar initiative in the 50 US states – a Geoscience Information

Network. These and other continental initiatives will be well linked to ensure complementarity of development and maximum synergy and benefit globally.

7. CONCLUSIONS

This paper has presented how Geological surveys, which exist in one form or other in every European Member State, are attempting to make their data more accessible; an objective that is central to their missions to acquire, analyse, manage and disseminate geoscience information and knowledge, in order to contribute to the quality of life and economy of their nations. When these surveys were founded (many in the mid-1800s), the factors driving geological surveys were scientific understanding and frontier exploration. These drivers have evolved over the years, through resource assessment and economic development to the present day concentration on environmental issues and societal benefit. In response most national geological surveys have changed the focus of their activities, from initial primary mapping, through resurveys, to thematic mapping and the majority are now embracing client-driven research and an increasing number (though unfortunately far from all) have given priority to the improved management and digital dissemination of their unique data and knowledge bases.

Despite most surveys having over 100 years of public service and geological map production behind them, the relevance of geology to managing our environment (both resources and hazards) is still not as widely understood outside the geoscience profession as it should be. Regrettably, too many of today's decision-makers (and fund-holders) – the politicians and the businessmen - are often woefully unaware of the critical role geological factors, and thus geological spatial data, can play in disaster mitigation and planning, environmental protection and sustainable resource development. This means that the comprehensive and unique information and knowledge bases that exist within each national survey remain, sadly, under-exploited.

Geological surveys have to accept a significant share of the responsibility for this. While there is an absolute necessity for a strong foundation of high quality geoscience research there is also an urgent need to re-balance an outmoded focus on traditional complex and narrow "academic" geological map output. To make sure geological knowledge plays a full role in the development of society geological surveys need to build products and services that ensure their data is widely accessible and genuinely meet society's requirements. To do that these products and services need to be *accessible* in a way that is meaningful to the largely non-geoscientific end-user and made available to them when they need it and in the format they require. Today that means web services. Inexpensive GI and web technology means all geological surveys have the potential to deliver flexible products and services from their unique geoscience knowledge bases.

OneGeology-Europe is seeking to help make this happen, to improve the use and exploitation of geoscience data and to ensure that data plays a full part in a future European environmental spatial data infrastructure.

ACKNOWLEDGMENTS

In addition to thanking the European Commission for funding OneGeology-Europe, I would like to acknowledge the work of scientists, informatics experts and managers in geological surveys and global and regional organisations around the world. Without their effort, commitment and goodwill OneGeology and OneGeology-Europe would still be concepts.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Quality of Life Environmental Information Services: a SEIS forerunner

K. Karatzas

Dept. of Mechanical Engineering, Informatics Applications and Systems Group, Aristotle University, GR-54124 Thessaloniki, Greece (kkara@eng.auth.gr)

Abstract: The Shared Environmental Information Space (SEIS) is a concept based on the Shared Environmental Information System and the Single European Information Space, and aims at the seamless access, exchange and reuse of environmental information towards the improvement of environmental decision making and quality of life, as well as the support of e-governance and business. The SEIS concept may be promoted and served with the aid of environmental information services of the digital era, that make use of user friendly ICT and human-centric knowledge engineering for the advancement of the quality of life for the European citizen and the transparency towards and access to information concerning the environmental domain. The present paper discusses the linkage between SEIS and Quality of Life (QoL) information services, and provides with an overview of the current status and future developments in this field. On this basis, it is suggested that QoL information services may serve as a paradigm, for the implementation of the SEIS and its deployments towards public administration and the citizens.

Keywords: Shared Environmental Information Space, Quality of life, Electronic Information Services

1. INTRODUCTION

The Shared Environmental Information Space (SEIS) for Europe may be considered as a vision and a wishful thought, in line with the need for seamless access to information. It is a concept in the hart of the concept for environmental information access anytime, anywhere, for anyone, in an easily to be used way, towards serving the most vulnerable parts of the population. SEIS may be placed in the core of infrastructures and arrangements that need to be materialised in order for Europe to move towards the future in a modern, human centric and more efficient (from the administrative point of view) society. Nevertheless, the SEIS is also a reference domain for future information services that optimise the reuse of public data – public sector information, integrate multiple scales of scientific research, support the implementation of new environmental legislation, and thus maximise the profit for the citizen, in terms of quality of life.

Such services address some “obvious” questions: which are the environmental pressures and the boundary conditions that frame the urban web within which the citizen has to live in and operate? And how do they interlink with contemporary digital cities and perceptions of digital services in the e-participation society?

Among environmental quality domains, air quality (AQ) is among the most significant and well studied areas. Yet the question of perception, interpretation and communication of AQ information remains open, and needs to be addressed in an effective way in order to design better early warning and information systems for the protection of human health and new information services for quality of life support. Contemporary AQ information dissemination methods and tools can now make use of various telecommunication channels for pull and push service provision. Mobile devices like phones, PDAs, in combination with mobile internet and broadband technologies may help in the design and personalisation of early warning, location-based, quality of life services. Moreover, communication per se is not based solely on written or oral language forms, but makes use of graphical, symbolical and in more general terms multimedia language communication schemes, via properly designed multimedia environmental information content, streamed via multimodal communication channels. Human perception and communication of such information is based on design schemes that may convey the message in an understandable and effective way.

The present paper provides with information on atmospheric environment (air quality and pollen related) information services (present and future), and on the way that such systems may support the quality of everyday life. This discussion is done in order to serve the need for a SEIS service paradigm, which may be used as a reference as well as a lever for the practical implementation of SEIS technologies and concepts.

2. THE SEIS CONCEPT

The production of environmental information in a digital manner has been included in a number of EU directives for some years now. Maybe the most pronounced environmental thematic area in this respect is air pollution, where member states had to “produce” via monitoring (and modelling) air quality concentration estimations, on the basis of predefined monitoring technologies and protocols, and communicate them to the EEA via the EIONET network. EIONET is a partnership network of the European Environment Agency (EEA) with countries that are either members or co-operators, and supports the collection and organisation of data and the development and dissemination of information concerning Europe’s environment (http://www.eionet.europa.eu/) EIONET was developed with the support of IDA, which is the EU programme for the interchange of data between administrators. IDA is now part of IDABC, which stands for Interoperable Delivery of European eGovernment Services to public Administrations, Businesses and Citizens (http://ec.europa.eu/enterprise/ida/index_en.htm). EIONET (already available by the end of the 90ties) provided with a physical network that supported the connection of EU member states, and promoted the communication of environmental information.

As environmental information production increased in the last decade, it became evident that such information represents a large portion of the so called Public Sector Information, i.e. information produced by or for public bodies, for serving operational and –or legal mandates. The accelerating maturation of e-services in both the public sector and mostly in the ICT industry, together with the developments in technology (like the semantic web, the web 2.0 and the post 3G mobile phones), prioritized the aspect of the decentralised use of environmental information, the ease of its re-use, the interoperability and cross accessibility of services and tools, and the potential of environmental information to support citizens and decision makers in their needs. The concept of SEIS developed within these lines of thought, with the aim to simplify the collection, exchange and use of the data and information required for the design and implementation of services. Such services will have a direct reference to the everyday life, will be framed and formulated by the users on the basis of profiling and georeferencing technologies, and will also support the implementation of environmental policies. The SEIS concept is being developed in parallel with the concept of the Single European Information Space, which is the first objective of the EU policy framework for the information society and media (i2010).

The principles of the SEIS are related to the ones of the Shared Environmental Information System, as stated in a relater document of the EU(COMM(2008) 46 final)[1] and are presented in parallel with the latter in the following table:

[1] Brussels, 1.2.2008 COM(2008) 46 final. Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions: Towards a Shared Environmental Information System (SEIS).

Table 1. SIES and the Shared Environmental Information space concept comparison.

Shared Env. Information System (SEIS)	Shared Env. Information Space
Information should be managed as close as possible to its source	Information should be made available for reuse as close as possible to its source
Information should be collected once, and shared with others for many purposes	Information should be collected from various sources, and be placed in the same information pull, to be reused
Information should be readily available to public authorities and enable them to easily fulfill their legal reporting obligations;	Information should be readily available for content and service providers as part of digital services
Information should be readily accessible to end-users, primarily public authorities at all levels from local to European, to enable them to assess in a timely fashion the state of the environment and the effectiveness of their policies, and to design new policy;	Information may originate from citizens, coming from citizens, in the form of on the fly measurements and observations. Such information should be made available freely to all interested parties, following the concept of social media, and under a predefined technical format.
Information should also be accessible to enable end-users, both public authorities and citizens, to make comparisons at the appropriate geographical scale (e.g. countries, cities, catchment areas) and to participate meaningfully in the development and implementation of environmental policy	QoL information services should be made available, allowing for everyday life pattern formulation and comparison between European cities and citizens, on the basis of predefined indicators.
Information should be fully available to the general public, after due consideration of the appropriate level of aggregation and subject to appropriate confidentiality constraints, in the relevant national language(s)	Env. Information should be made available to the public under the same time frame as the env. problem described, i.e: AQ episodes in terms of the time frame of an hour, AQ strategies in terms of weeks to months, etc.
Information sharing and processing should be supported through common, free open source	A reference set of digital services should be made available for free from European Environmental Authorities, accompanied by user scenarios, and be ready to be included in commercial information portals and mobile phone-device applications. These services should be accompanied by a basic toolbox, and both should be provided as open source software resources.

SEIS poses the same challenges towards its adaptation and implementation as the Shared Environmental Information System. Thus, there is no implementation guidance concerning the integration of the SEIS in the relevant legislation on EU level. In addition, harmonisation and standardisation activities are still lacking. Moreover, tools for the support of the SEIS like taxonomies, ontologies and services have not emerged in a usable form so far. In addition interoperability between environmental domains is still an open issue. Overall, what seems to be missing is a well targeted, easily recognisable and understandable paradigm of a service that promotes the concept of the SEIS. Such a service should prioritise the citizen in order to gain social visibility and support, and support environmental information provision, better environmental policy making, and ultimately better quality of life. On this basis, QoL environmental information services can be used as the reference-paradigm for the development and the implementation of SEIS services in Europe.

3. ASSESS TO ENV. INFORMATION CONCEPT AND AIR QUALITY

Environmental Information corresponds to the awareness that a society poses in terms of the quality of the environment and its vision towards the improvement of everyday life. Nevertheless, in the early days of environmental practice, information about the quality of the environment was considered to be equivalent to classified governmental data and thus should not be revealed to citizens without certain limitations being applied. As public participation and policy making developed rapidly in Europe, it was recognized that citizens well informed in environmental issues can support the formulation and application of practises for the protection of the environment and the support of sustainable development (a development that optimises the today's usage of natural resources on the basis of environmental criteria for tomorrow). Yet, what was lacking (and it still is), is a model for effective communication of environmental information to the public. This type of model – guidelines, started to rise within the related legal framework, that, in its turn, tried to incorporate in the best way the scientific knowledge concerning the impacts of AQ to human health and the environment, and introduced in our everyday vocabulary terms like assessment, limit values, target values, concentration and many others.

The first EU legislation concerning air quality information availability was Dir. 82/459 later replaced by Decision 97/101 which stated that E.I. should be made accessible to the public via an information system set up by the European Environment Agency (EEA). This decision led the EEA to establish the European air quality information system Airbase. Airbase provided with a backbone for information communication but, most importantly, with a platform for the development of the idea that EI should be made available to the public via automatically operated communication channels.

A major change came with Directive on Ambient Air Quality Assessment and Management (96/62/EC), which required for the development of action plans concerning zones within which concentrations of pollutants in ambient air exceed limit values. These limit values were established by new (Daughter) Directives, that replaced old ones. It is worth noting that within these Daughter Directives, the use of computer-network services is mentioned in order to provide the public with the appropriate air quality information, which should be up-to-date and should be routinely made available (Karatzas and Moussiopoulos, 2000).

On the other hand, the EU legislation concerning public access to environmental information developed already in 1990 with directive 90/313/EEC. This directive was amended by the "Directive of the European Parliament and of the Council on public access to environmental information", that later became Dir. 2003/4/EC, which declares that environmental information should be provided to the public on-line. This specific request results from Article 7 of Directive 2003/4 which states that "Member States shall take the necessary measures to ensure that public authorities organise the environmental information… with a view to its active and systematic dissemination to the public, in particular by means of computer telecommunication and/or electronic technology".

The most important step towards a SEIS for Europe in the area of air pollution, as well as for the electronic information services for the environment, was done with Dir. 2008/50/EC. This directive (called CAFÉ or Clean Air for Europe Directive), specifically states that the public should receive information for the ambient concentration of a number of pollutants including sulphur dioxide, nitrogen dioxide, particulate matter (at least PM10), ozone and carbon monoxide (Dir. 2008/50/EC), thus posing the issue of seamless access to this information all over Europe. According to the same legislation, this information shall be updated at least daily, and when practicable on a hourly basis, and should be provided timely or even in advance, on the basis of forecasts. It should be noted that the same directive makes a clear reference to the use of the internet and any available telecommunication means, and shall also take into account the provisions of the INSPIRE directive 2007/2/EC. On this basis, the citizen has right to be informed for any type of exceedance, its location, duration, and impacts to the quality of his/her life, both timely and in advance. This legal mandate brings into the first priority of action QoL information services, and may be used as a reference for all other environmental domains.

4. QUALITY OF LIFE INFORMATION SERVICES

4.1 The outriders

There is a variety of environmental quality information providing/disseminating paths, all being directly connected to the operation of "traditional" and electronic communication schemes. One of the most "complete" examples is that of the air quality information services that may be found in the results of the APNEE and APNEE-TU projects (2000-2004, www.apnee.org). APNEE established a multi-channel information service platform for the dissemination of air quality information (and in this way provided with a platform for the implementation of e-government information services), via using appropriate information presentation methods. The dissemination platform made use of various telecommunication channels for pull and push service provision, indulging internet for e-mail notification, world wide web for detailed pollution related information, SMS for early warning services, WAP and J2ME applications in mobiles and PDAs for enhanced graphical and informative content on the move, street panels (VMS) for covering key parts of the urban web and voice services for personal communication support. Moreover, APNEE provided with location based services, and supported personalization of the information channel(s), content, and frequency of information services and related alerts.). In addition, some other projects like Marquis (2005-6, http://www.marquisproject.net/) included in their approach the concept of automatic multilingual text generation for the support of environmental information services. Moreover, a number of projects and R&D initiatives have been investigating and developing air quality information services (Westbomke et. al, 2003 ; Karatzas et. al., 2004 ; Arauco and Sommaruga, 2004 ; Dickinger et. al., 2005 ; Storch, 2005. Some of those projects are AIRTHESS (www.airthess.gr); Luftkvalitet (www.luftkvalitet.info); YourAir (www.airtext.info); CITEAIR (citeair.rec.org), AIRALERT (www.sussex-air.net), and 191ASK (www.breathebetter.com.au).

In terms of environmental information dissemination technologies, system modules should be adaptable to content needs and dissemination requirements via component oriented architecture. The system will "materialise" environmental information services on the basis of "smart" spatio-temporal profiling of the knowledge base. The latter may consist of sensor data and model results, on top of which a semantic layer can be constructed so as to describe dependencies and usage scenarios on a level that may directly be addressed by information dissemination channels.

4.2 The forerunners

Human activities in cities include commuting, working, and in house activities. Air quality related environmental pressures are thus always present, and for many categories of citizens like asthmatic, elderly or children, it is a major concern and a quality of life paragon. Thus, alerting on a personal basis is required, that addresses information needs tailored to living patterns of individuals, for multiple spatio-temporal scales. In addition, each citizen would prefet to define and assess quality of life indicators on a personal basis. A service approace may be used, focusing on mobile devices (phones, PDA's, WIFI available devices, etc), that will allow for personalisation of alerting messages, via multiple communication channels. In addition, a wiki for e-tags related to events and features of interest within a city may be created on the basis of the SEIS, accompanied by a blog-like function, tailored to be available via mobile devices. It is therefore expected that a folksonomy of environemntally aware, e-active citizens will be developed, that will provide with added value information and warnings to its members, while supporting self orgqnisation and equal participation of citizens to environemntal decision making in the city. In this way, the SEIS will not only be used, but also enriched with new information, in a way similar to the one that is applied in social media nowadays. The environment can thus be turned into the bussword of the day, and individuals will be able to define, develop and share their owninformation services and environmental resources. In this way the SEIS may become familiar and appealing to the end user, while environmental decision making will also be advanced.

4.3 QoL service examples

Human activities within the urban web (commuting, resource usage, services usage), may be supported by smart information services that will advance quality of life and lower environmental pressures. The new concept that such services introduce is that of quality of life information services, which may be tailored and personalized on the basis of individual citizen profile, as in the following examples:

- Commuting route within a city: Alternatives may be prioritized on the basis of availability, service robustness, health exposure risks for asthmatic persons, sustainability measure in terms of energy consumption, etc
- Health status: asthmatic people, sufferers from cardiovascular diseases, children, elderly etc will be able to tailor the way that they will receive health warning and advices, on the basis of geo-referenced information related to hotspots within the urban web, street canyon related air pollution levels, microclimatic conditions, etc.
- Information presentation and interpretation means: advanced 3-D visualization of the air pollution levels and multimedia means will be combined with simplified, well recognizable health risk alerts, while SMS and other types of simple text massaging may be used as a stand alone way of information communication, or in combination to a graphical way of information presentation and interpretation

Last but not least, it should be noted that up to now the citizen was a passive receiver of information and advices concerning the quality of the environment that he/she lives in. Emerging sensor technologies as well as the development of ICT are bringing in the scene sensors that will "emit" information for everyone to use, virtual electronic urban spaces that will support smart tagging and information posting and exchange, and mobile devices that will support social media and end-to-end citizen communication and information sharing. On this basis, the citizen will be able to take part in the monitoring of the urban environment quality, and "emit" this information to other citizens, with the aid of e-tags and by making used of on-the-fly social media networks, that may be posted to an urban environment information space which will be made available via the SEIS. This space may be envisaged as a thematic geospatial wiki that may be accessed via mobile devices, and to which citizens may post information, advices and warnings for others to read and react, in addition to the advices and warnings that they will receive from a central system. In this way, the citizens are expected to become more environmentally active and aware.

CONCLUSIONS

The emergence of a Shared Environmental Information Space will speed up the application of advanced, Quality of Life (QoL) information services that will allow for personalised, flexible content provision and communication mode (pull-push). The air quality domain may serve as a reference domain for the development of such services. Moreover, technological developments in micro-sensors and related ICT, will allow for the enrichment of information via multiple monitoring data coming from people commuting within a city, monitoring the environment, sharing their observations and suggestions, and thus creating a live and active participation society on the basis of QoL information services. Thus, QoL services may be used as a paradigm for the establishment and use of the SEIS in Europe.

ACKNOWLEDGEMENTS

This work is related to COST Action ES0602: Towards a European Network on Chemical Weather Forecasting and Information Systems, www.chemicalweather.eu.

REFERENCES

Arauco E. and Sommaruga L. (2004). Web Services for Environmental Informatics, iEMSS proceedings, pp. 543-549 (http://www.iemss.org/iemss2004/pdf/envinfo/arauwebs.pdf)

Dickinger A., Heinzmann P. and Murphy J. (2005), Mobile Environmental Applications, Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS'05) - Track 7, p. 178a http://csdl2.computer.org/comp/proceedings/hicss/2005/2268/07/22680178a.pdf

Karatzas K. and Moussiopoulos N. (2000) Urban air quality management and information systems in Europe: legal framework and information access. *Journal of Environmental Assessment Policy and Management* **2**, 263-272

Karatzas K., Nikolaou K. and Moussiopoulos N. (2004). Timely and valid air quality information: The APNEE-TU project, Fresenius Environmental Bulletin, Vol. 13, No.9, pp.874-878.

Storch H. (2005). Environmental Collaboration through Location-Based Mobile Messaging and Communication Tools, EnviroInfo2005, pp. 835-839

Westbomke J., Dombeck Th., Haase M. Ebel R. and Lehne D. (2003), WAP-Based Environmental Information Services, 17th Int. Conference: Informatics for Environmental Protection, Vol. 1., pp. 185-192. Metropolis Verlag, Marburg, Germany.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Reporting Environmental Information Best Practice and 15 years of experience in Germany

W. Kazakos, G. Lukacs[a]

[a] *disy Informationssysteme GmbH, 76133 Karlsruhe, Germany, (kazakos|lukacs@disy.net)*

Abstract: In the last 15 years many state and federal level environmental authorities in Germany have cooperated by joining their financial resources and competences to create new software solutions for environmental and spatial reporting – long before INSPIRE and SEIS. The cooperation was initiated in 1994 by Baden-Württemberg, a German state with an out-standing environmental information system. disy Informationssysteme GmbH is working in the related umbrella projects, among other companies. The paper tries to give a first overview of the platform and the challenges addressed. It can be seen as a best practice approach which might help to understand the technical complexity in developing SEIS.

Keywords: Environmental reporting, Environmental data; Environmental information; Environmental information space.

1. INTRODUCTION

In the last 15 years many state and federal level environmental authorities in Germany have cooperated by joining their financial resources and competences to create new software solutions for environmental and spatial reporting – long before INSPIRE and SEIS. The cooperation was initiated in 1994 by Baden-Württemberg, a German state with an out-standing environmental information system. disy Informationssysteme GmbH is working in the related umbrella projects, among other companies.

The know-how and the platform disy Cadenza for data integration, analysis and reporting is a result of the ideas and requirements developed by German environmental authorities. The initiative incorporates R&D and technical competence in information systems, geographic information systems, decision support, databases, human-computer interaction and software technology. The platform is currently used in a large number of information systems across numerous environmental authorities in Germany. This paper provides an overview of the challenges addressed, the platform itself and its possible contribution to SEIS.

2. MAIN CHALLENGES OF ENVIRONMENTAL REPORTING

2. 1 Amount of data

A very broad spectrum of relevant data is covered by integrated environmental information systems. The sheer amount of data is the first challenge which has to be addressed. The database schema of an environmental information system even of only one German state such as Baden-Württemberg contains up to several thousand database tables, not to mention the geographical layers. Extensions and modifications to the database schema occur regularly. Changes and new requirements about how the data should be queried and presented are even more frequent.

2. 2 Multidisciplinarity

Environmental legislation is getting more holistic, requiring an integrated view on many related issues. The Water Framework Directive experts on the application area for example have to get a much wider overview on the available data, e.g. experts working primarily on groundwater quality get in contact with water usage rights. Working or even just getting an overview on databases one is not familiar with is very difficult and time consuming. Therefore, usability issues become central, especially those related to querying and visualizing structured database data.

2. 3 Joining know-how in data modelling and the application domain

The core technology of data modelling is quite straight forward as it has been researched and applied for the last 30 and more years in database science. The major challenge however is still to apply the know-how to the environmental application domain. Setting up environmental reporting applications needs always cooperation between environmental domain experts, who try to formulate their needs and requirements, and the IT and database experts, who creates data models and applications that are extensible and flexible enough for fast changes in application domain requirements. This knowledge is however rarely publicly available.

2. 4 Missing standardisation

As the environmental application domain is per se heterogeneous and environmental databases are usually developed by small local companies, the level of standardisation or even of universal data models is quite reduced. This has both advantages and disadvantages, e.g., it allows more variety, considers more local specialities, but is inefficient in the large. We believe, however, that at the current level of standardisation the disadvantages outweigh the advantages by far.

In addition standardisation efforts in the geographic domain, like OGC-standards, cover only geodata which are just one minor part of the general data management discussion for environmental data.

3. COOPERATION

To tackle these and other challenges and to reduce development cost a cooperation was initiated in 1994 by the Ministry of Environment and Transportation in the German state of Baden-Württemberg which comprises today more than 30 members and has become very successful in all parts of Germany. The cooperation includes German authorities as well as leading universities, research institutes and private companies. Through this, a large potential for innovations – essential for non-traditional tasks typical in environmental protection – combined with industrial strength and pragmatism has been present in the cooperation, representing another cornerstone beside innovations for successfully mastering the challenges.

The requirements on information systems in environmental protection, traffic and public administration resemble each other to a large extent in all German states and also at the federal level due to their common obligations. Therefore, a special cooperation in the form of an agreement called KoopUIS (Kooperation von Bund und Ländern für Konzeptionen und Entwicklungen von Software für Umweltinformationsysteme – Cooperation on Developing Concepts and Software Systems for Environmental Protection) has been initiated. In the meantime, all state-level environmental authorities and many federal agencies participate in KoopUIS. In the executive committee all state and federal level organisations are represented. The costs and the copyrights of the concepts and systems developed in a project are shared between the partners who participate in the specific project. Nevertheless, all projects are documented in the KoopUIS. Workshops and other meetings are organised in order to support the work and to gain new partners.

Several related R&D umbrella projects have been initiated by Baden-Württemberg's Ministry for the Environment in Stuttgart, Germany. The goal of these projects is to develop innovative, economically advantageous and powerful applications by bundling financial and human resources. The developed concepts and systems can be reused by all project partners. The assignment of the projects and the financing are agreed on separately between the contracting bodies and the contractors of each project. The contracting bodies coordinate their work in KoopUIS. The umbrella projects have a long history: the project GLOBUS (Globale Umweltsachdaten – global environmental data) ran between 1994-1999 and the project AJA (Anwendung JAVA-basierter und anderer leistungsfähiger Lösungen in den Bereichen Umwelt, Verkehr und Verwaltung – Application of Java-based and other Powerful Technologies in Environment Protection, Transport and Public Administration) between 2000-2004. The current project KEWA (Kooperative Entwicklung wirtschaftlicher Anwendungen für Umwelt und Verkehr in neuen Verwaltungsstrukturen - Cooperative Development of Economically Advantageous Software Solutions for Environment Protection, Traffic in Restructured Public Administration) started in 2005.

4. GERMAN ENVIRONMENTAL REPORTING PLATTFORM

One of the major achievements of the cooperation was the development of a common software platform, disy Cadenza, tailored to the needs of environmental reporting obligations in Germany.

The main driving force to develop the platform was the understanding, that analysing complex data, such as environmental data, requires powerful methods to select, prepare, analyse, visualise and finally report data – implementation of special purpose applications was technically and economically not anymore feasible.

The result is a configurable system general enough to offer the means for many analysis and reporting tasks for environmental authorities. The configuration is simple enough to allow experienced users to cope with the configuration. Less experienced users can use the information prepared by the experienced users.

4. 1 Data management strategy

As it is not realistic to standardize all data management and integration aspects for the environmental application domain immediately, the platform follows a twofold strategy in data management. On the one hand it allows the querying of all sorts of available heterogeneous databases, even if the database schema does not consider modern design principles. The person in charge can use wizards to connect to several data bases and to map tables and rows to those information views the user is expecting. Through web services the platform provides access to the data for third party applications such as web portals.

In parallel data management standardisation had been started to allow application of data model patterns, conventions for newly developed or reengineered of database schemata. It is important to note that by standardisation on the data base level the complexity can be reduced to a great deal, while standardisation on web service level solves only one minor part of the problem.

4. 2 Explorative querying, dynamic query feedback

Searching for data in a database the user is not familiar with is a frustrating and time consuming task. Explorative data analysis (Schneiderman/Plaisant 2005) gives the answer to this challenge. Processing and visualising topical, temporal and geographical data in the form of business diagrams, maps and map-based diagrams is the next challenge, requiring innovative solutions in user interface design and functionality. The central idea of the platform was derived from classical business reporting applications. However, experience showed that they had to be strongly adapted for environmental reporting obligations.

In form-fill-in querying the user has to work “blind”, i.e. formulate queries by filling in the query conditions in forms. To be able to do so, he or she must either know the data well or must make many trials to get a sense out of the data and to be able to formulate meaningful queries, i.e. queries returning neither too few nor too many results. In explorative querying – as opposed to form-fill-in querying – the user gets continuous feedback on the data. E.g., the application offers the available values for selection or gives a draft overview on the distribution of the data. Figure 1, (D) and (C) gives and example for temporal query condition which show a histogram of the available measures.

Explorative querying has proved by distance superior to form-fill-in querying, regarding both task completion time and the number of required queries to complete a task resulting in higher levels of user satisfaction.

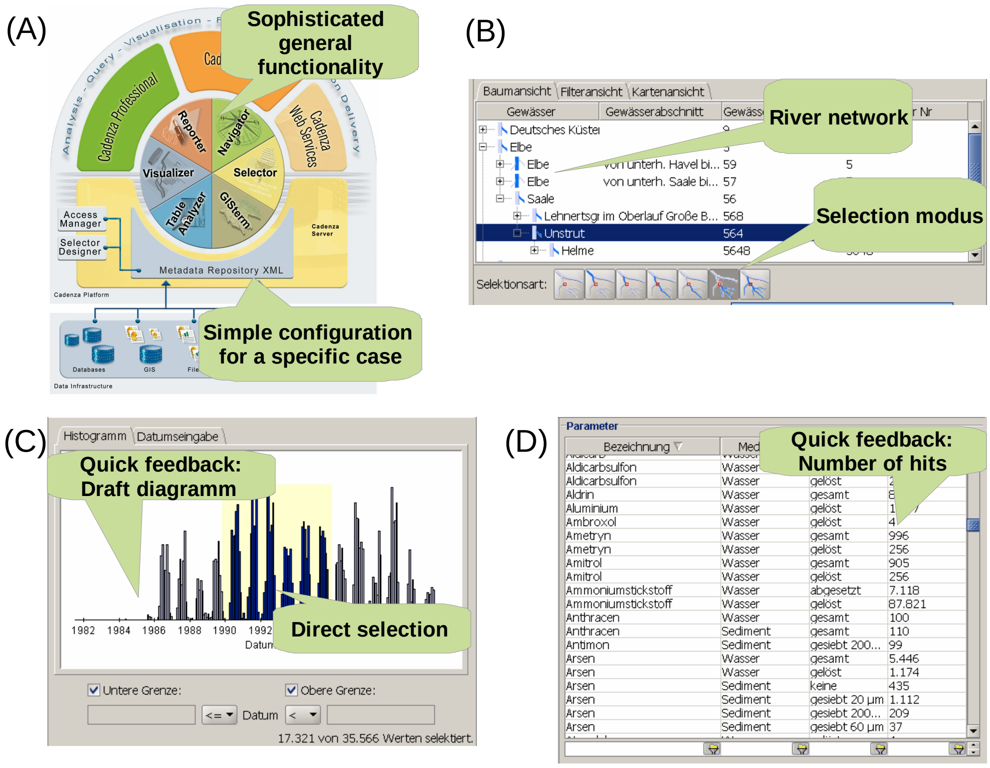


Figure1: (A) Architecture; (B) Querying with semantic structures (e.g. river network); (C), (D): Explorative querying with quick feedback and direct selection (e.g., temporal and topical query conditions)

“Overview first, details on demand”, or zooming, is another related requirement: The user at first wants to get an overview of the data and then zoom in on a particular (geographic, temporal or topical) “area”.

The platform incorporates such approaches and thus allows the user to browse and explore the data and to smoothly drill-down to the information he or she needs. The main difficulty was to incorporate spatial and non spatial criteria for the adaptive content selection, as most environmental information have regional (spatial) and thematical aspects.

4. 3 Semantic structures: River network, taxonomy, semantic network

Another issue was querying with some sort of semantic structure. An example is the task of finding all constructions in a river network that potentially hinder the migration of fishes (from the sea) to a specific point in a specific river (Figure 1, (B)). Ideally the user can formulate his or her information need very simply and the system takes over the potentially complex evaluation of the query.

Additional taxonomies like querying with class or order, or data indexed on the level of species can be used. The extension to ontology based query mechanisms is envisioned but far from been standardized.

4. 4 Chaining

In many complex issues, several query and analysis steps, one built on the other, have to be completed in a sequence; i.e. the result set of one query or analysis step has to be considered in the next step e.g. as a query condition. An example is to select groundwater quality values above a certain limit, then to select the water protection areas where the values are measured, and finally select the relevant water works.

The functionality of chaining allows exactly this: Along all (foreign-key, spatial or temporal) relations in the database the user can easily combine data over information views. Chaining is another essential feature of the platform, allowing to find the answer to a large number of simple or more complex information needs by easily combining several simple steps. Note, that without chains, information views should be prepared for all sorts of combined information needs that would not be possible in a somewhat more complex database.

The new community project – Linking Open Data (Bizer et al., 2008) – of the World-Wide Web Consortium, Semantic Web Education and Outreach Interest Group, follows an idea similar in many respects to the chains in the platform.

4. 5 General code and configuration: Executable object definitions

The platform must be independent of the specific database-schema, the specific data processing steps and the specific visualisations in a specific application, otherwise the development and the maintenance of the system would be prohibitively expensive (e.g., the environmental database of Baden-Württemberg covers several thousand database tables and there are about a thousand different information views on them). One major requirement was, that the platform had to offer a general infrastructure that can be easily configured to the needs of a specific application. disy Cadenza is built up in this way: the general functionality can be configured by a metadata repository to the needs of a specific application. Figure 1, (A) provides an general overview of the approach, including the connection to the different data bases through a metadata repository, the access manager and designer tools to define these connections as well as the three major access methods, i.e. through a desktop application, a web application or through web services.

Through splitting the sophisticated functionality and its use for a specific application, technical complexity can be reduced to a large extent. Furthermore, sophisticated skills are only needed in developing the general application, but not for its use for a specific application. Therefore, a relatively small developer team with sophisticated skills can produce a reporting application which then can be configured by a much larger number of people, having more background in the application area than in software development.

5. CONCLUSIONS FOR SEIS

disy Cadenza as a platform incorporates more then 15 years of experience in environmental reporting in Germany and several hundred person years in development. This paper mentions just some major challenges and approaches to address them. The challenge of implementing the Inspire directive on a European level, which can only be seen as one part of SEIS, shows already the complexity which will even increase by far when addressing more than geographic information.

Drill-down information on a European level and reporting will need new forms of cooperation, standardisation and applications. The German example shows a way how to address these issues. Standards on Web Service level are needed and important, but blind reliance on web services will not be sufficient, as data amount and heterogeneity cannot be fully solved without classical data management and integration approaches.

REFERENCES

Bizer, C., Heath, T., Berners-Lee, T.: Linked Data: Principles and State of the Art. 17th World Wide Web Conference (WWW2008), W3C track, Beijing, China, April 2008.

Briesen, M., (et al.) (2006) Information System for Implementing the European Directive on Environmental Noise in Bavaria In: Managing Environmental Knowledge (EnviroInfo 2006), Tochtermann/ Scharl, Shaker Verlag, 2006, S.525ff

European Commission (2008): Action plan on sustainable production and consumption and sustainable industrial policy.

Kazakos, W. (2006): Metadata Management with ISO 19115 – The Key to Success of Spatial Data Infrastructures (SDI). In: Managing Environmental Knowledge (EnviroInfo 2006), Tochtermann/ Scharl, Shaker Verlag, 2006, S.257ff

Lukács G. (2003a) Decision Support for Consumers – Information Services for a More Sustainable World. In A. Gnauck an R. Heinrich (editors): Proc of the 17th International Conference Informatics for Environmental Protection, Metropolis Verlag, Marburg,

Lukacs, G. (et al.) (2003): Data Warehouse for Water Management – Geographic Data and ETL Proc-esses. 17th International Conference on Informatics for Environmental Protection, Cottbus.

Lukacs G.: DataWeb for Integrated Environmental Information Systems - An Abstract View on disy Cadenza In A. Möller et al. (editors): Environmental Informatics and Industrial Ecology. 22th International Conference on Informatics for Environmental Protection, 10-12 Sep. 2008, Lüneburg, Germany, pp. 555-558, Shaker Verlag, Aachen, 2008

Lukacs G. (2008) Potential in Information and Communication Technologies for Supporting Sustainable Consumption. In A. Möller et al. (editors): Environmental Informatics and Industria Ecology. 22th International Conference on Informatics for Environmental Protection, 10-12 Sep. 2008, Lüneburg, Germany, pp.409-412, Shaker Verlag, Aachen.

Mayer-Föll, R. (et al.) (2004): Projekt AJA, Anwendung JAVA-basierter und anderer leistungsfähiger Lösungen in den Bereichen Umwelt, Verkehr und Verwaltung, Phase V, Forschungszentrum Karlsru-he, Wissenschaftliche Berichte FZKA-7077.

Mayer-Föll, R. (et al.) (2005): Integrated Analysis of Environmental Data through Networking and Innovations -- Despite Organisational Changes, Increasing Requirements and Decreasing Budgets. 19th International Conference Informatics for Environmental Protection, 2005.

Meinecke, S. (et al.) (2005): An information system for accounting fate of man-induced substances and determining their effects on aquatic organisms and ecosystem functions. 19th International Conference Informatics for Environmental Protection, 2005.

Schneiderman, B., Plaisant, C. (2005): Designing the User Interface. Pearson, Addison Wesley.

Wiesel, J. et al. (1994): GLOBUS - Konzeption und prototypische Realisierung einer aktiven Auskunftskomponente für globale Umwelt-Sachdaten. Abschlußbericht Phase I. Umweltministerium Baden-Württemberg

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Integration of national sources for the implementation of SEIS in Slovakia

Ing. Ján Tóbik., Ing. Martin Koška, Ing. Ján Cimerman

Slovak Environmental Agency (SEA) , Banská Bystrica

Abstract: The SEA is responsible for the management of metadata information system, which covers the metadata of the whole branch and integrates metadata from information systems, which are operated by the Ministry of Environment of the environment of the Slovak Republic. This solution leads to the unification of information systems at metadata level. The user thus obtains a consistent approach to descriptive information relating to the heterogeneous information systems.

Another step towards integration is the incorporation (implementation) of OGC standards into existing metadata information system. This process was initiated by the adoption of the INSPIRE Directive. The metadata information system obtained the ability to search for metadata records in the other European metadata information standardized systems.

Keywords: SEA, Metadata; OGC standards; INSPIRE Directive; Integration.

1. INTRODUCTION

The massive production of data and information from various, mostly heterogeneous resources in the environment area, moreover, accelerated by development of the Internet did not brought automatically better information to consumers. Even before more than 10 years ago started to work the first solution that integrated the majority of the national resources in the environment area into one virtual information system so that individual elements of the system maintain their independence and their own infrastructure. With the support of Slovak legislature have been gradually introduced into operation following registries:

- The catalogue of data sources (KDZ).
- Specific publicly accessible listing (OVPZ).
- EnviroInfo - Metadata information system.

2. THE CATALOGUE OF DATA SOURCES (KDZ)

Following up the efforts of the previous period after the introduction metadata information system into practice there was established a "The catalogue of data sources" as the first metadata information system in the environment branch. Since then, has elapsed 7 years now. "The catalogue of data sources" (KDZ) was created as index in the year 2001 and operated in terms of the general requirement of collection and disclosure of metadata connected to the information created or obtained the Ministry of Environment Republic of Slovakia and the Ministry of environmental organizations. Metadata should be accessible through the electronic catalog available via the Internet so as to serve a wide professional and laicus public.
„At the origin KDZ one of the objectives was to create the system that would comply with the existing and upcoming standards defining the structure of metadata information systems to ensure harmonization with similar metadata information systems at home and abroad.“[1]

3. THE SPECIFIC PUBLICLY ACCESSIBLE LISTING (OVPZ)

The issue of metadata is dealt with by the international standard ISO 19115, and for an understanding the efforts to establish the system for the collection of metadata (metainformation) is necessary to define the concept. METADATA term is generally designates "data about data or structured data about data." Metadata is essentially equivalent to the bibliographic indexed record. [2]

By approvement of Act no. 205/2004 Zz, which governs the conditions and procedure for collection, storage and dissemination of information on the environment, it was created space for the application of the requirements arising from this law, and thus allow access to information relating to the environment for public authorities and other legal bodies and individuals designated by this Act. Index "The specific publicly accessible listing" (OVPZ) has operated since 2005.

The operation of KDZ and OVPZ ensured the application of the laws into practice and disclosure of environmental information the general public.

4. ENVIROINFO – METADATA INFORMATION SYSTEM

After admiting of Slovakia to the European Union in 2004, our country accepted the pledge to adopt generally binding regulations of the European Union to implement them into the national structures. The legislative process also affected the area of metadata.
In the year 2004 the European Commission formulated a proposal of directive for the European Parliament and Council on the establishment of the Infrastructure for Spatial Information in the European Community (INSPIRE). The ambition on the creation of a standardized data collection was based on the requirements to ensure a uniform structure of the data provided, and thus contribute to the potential utilization of data throughout the European Community.
INSPIRE Directive 2007/2/EC, not only defines the metadata, but also exactly sets out the timetable for their creation. From 15.5. 2008 is in force implementing Directive for metadata. Its text as well as the directive itself is available at the address: http://www.sazp.sk/inspire/
Following the above, there was a mergence of a KDZ and OVPZ into a single department metadata information system of environment branch under the title EnviroInfo. It is available at the address: http://enviroinfo.enviroportal.sk.

"The Metainformation system is one of the key elements necessary to create infrastructure for spatial information in terms of INSPIRE. This framework provides a link to national and international level. Also issues of standardization and harmonization were taken into account, one the first aspects of its implementation has been dealing with was the design and implementation of metadata profile in terms of ISO 19115". (http://enviroinfo.enviroportal.sk/article/1)

Following the adoption of the INSPIRE Implementing Rules for metadata will be necessary to review the structure of EnviroInfo and to adapt it to the structure, which is based on European legislation. This step should ensure that EnviroInfo system will be interoperabile with other metadata information systems operating by European Union Member States.

This step will enable the user to search for the metadata, not only connected to the Slovak Republic. The applicant will thus gain an overview of the documents, databases, vector and raster GIS layers or maps already created and services offered across the European Union.

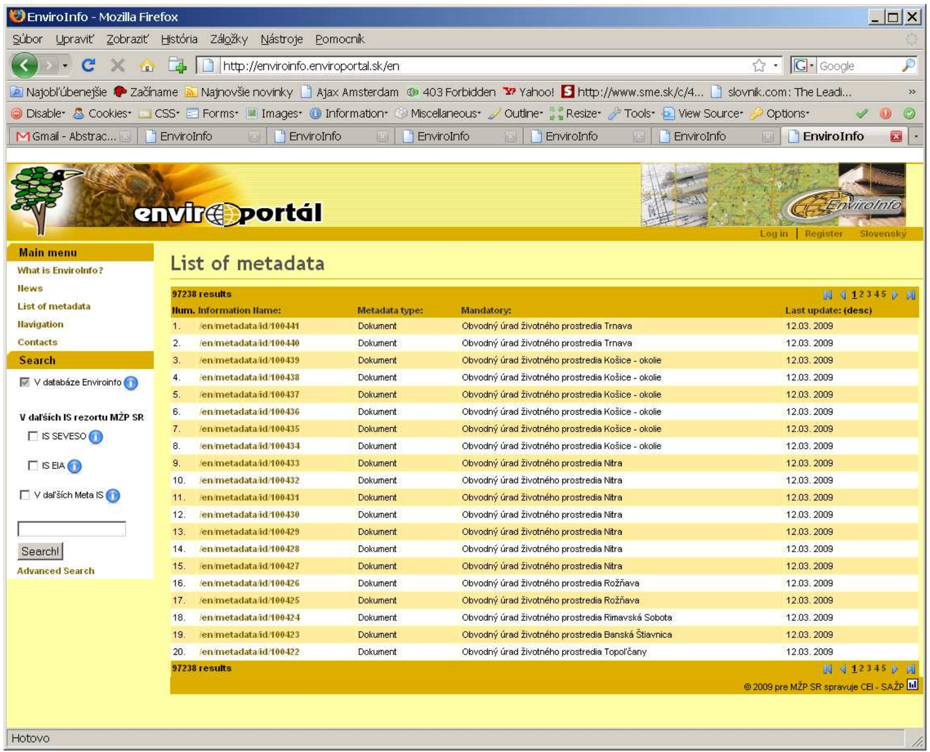


Figure 1: List of metadata records filled up by registered users

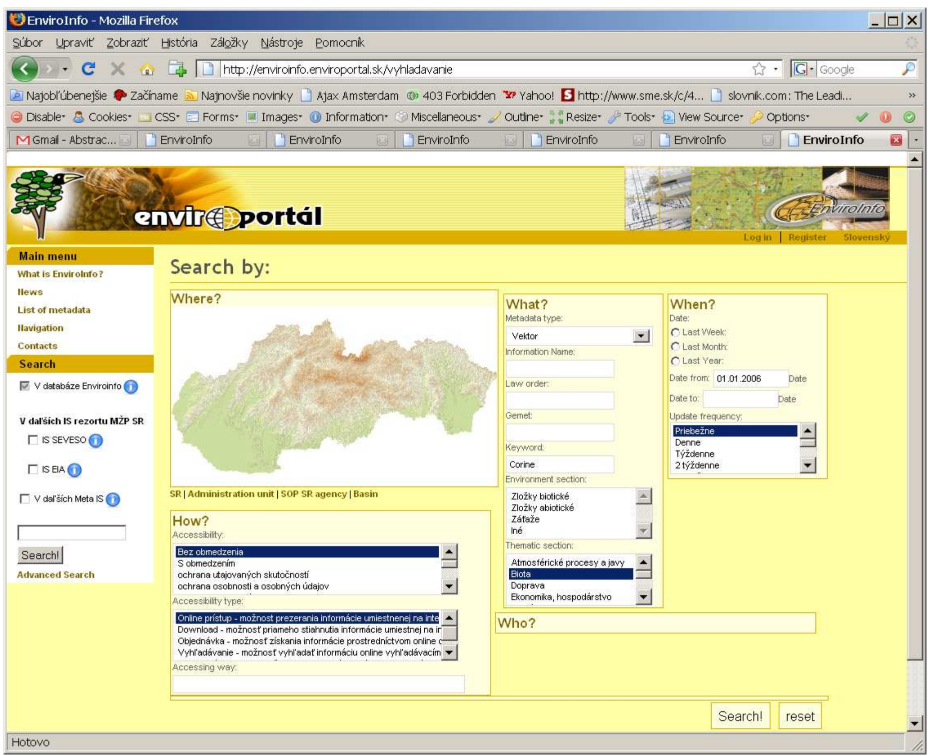


Figure 2: Modul for advanced search of particular metadata

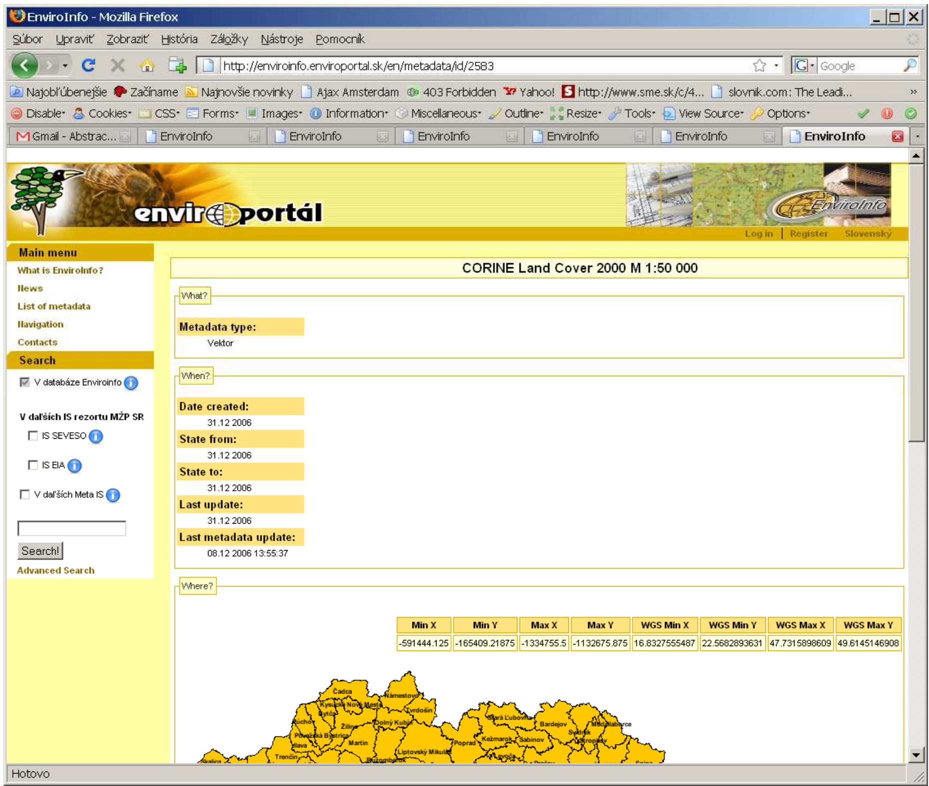


Figure 3: Result searched out record of particular metadata – upper part

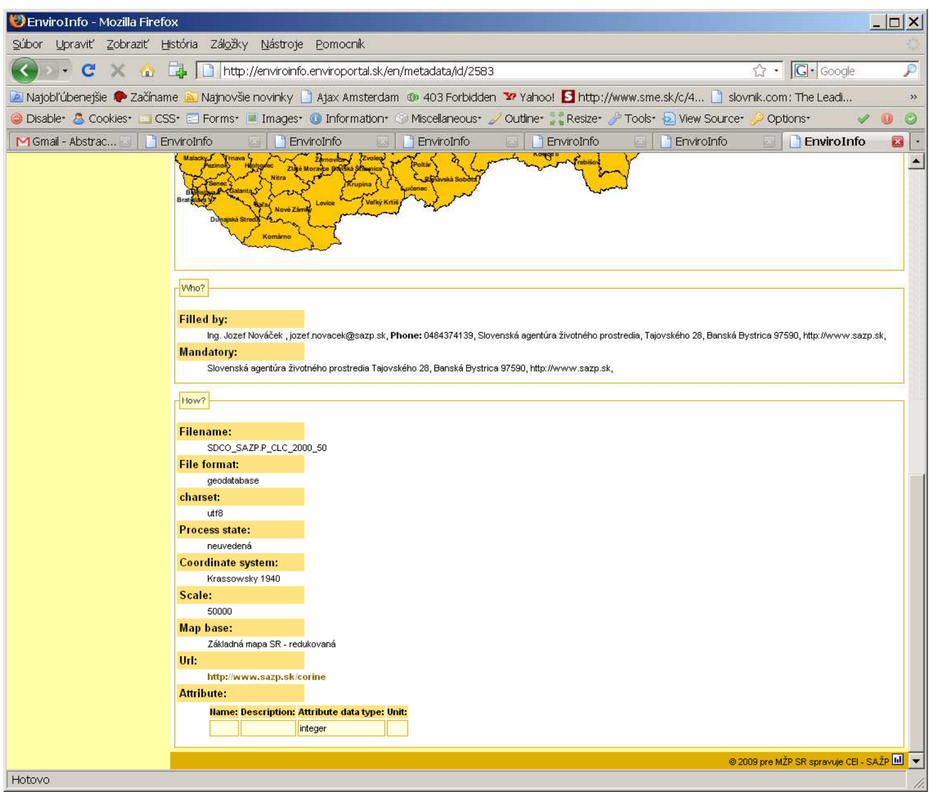


Figure 4: Result searched out record of particular metadata – lower part

Metadata information system is created as a separate web application in accordance with the paradigm three tier architecture. The application, presentation and database layer are treated as independent and interchangeable, ensuring modularity and portability of the system. In the application layer is solved the communication with web services, which represent a further source of data. In a further development envisages increased use of Web services, simplifying the user interface and powerful search mechanism.

Access to the editorial section EnviroInfo system is provided on the basis of registration and authorization. Authorization of users is preceded by checking by the administrator to list of charged and responsible persons who have granted authorization to enter into the editorial section. Thus, at the same time ensuring the protection from abuse of metadata information system.

The major part of the metadata documents were issued by authorities of the environment officies, but in the same way the user gets the other data, for which EnviroInfo system was created. List of charged and responsible persons is available at EnviroInfo site. (http://enviroinfo.enviroportal.sk/article/3)
By now it is possible to enter into EnvitoInfo metadata for the following items: a document database, map, vector, raster, others. Preparing the possibility of entering the metadata for the services. This structure is in terms of Act no. 211/2000 on free access to information and amending certain laws meets duty of free access to information. Operation EnviroInfo facilitates access of citizens to information, and also has an impact on the officials, because the information the applicant has the option to come up with a specific requirement, or the whole deal on access to information restrict to the area of the Internet.
EnviroInfo is designed as a tool for creating and searching metadata. If a user of EnviroInfo knows what he/she is looking for, eventually if for some reason he/she is interested in a particular part of the territory, just EnviroInfo is a tool that will provide a comprehensive overview of the data from the environment.

5. CONCLUSIONS

The predecessors of EnviroInfo - KDZ and OVPZ had ambitions to offer easy and scalable birds eye view at country of enviromental information. Utilizing their pros and cons continue the developers of EnviroInfo. The idea here is to cover all information relating to the environment by the national metadata information system and link it with the other metadata information systems so that the user could obtain by simple way the information that is important for further decisions.

The process of development of EnviroInfo is based on the latest legislation, developers themselves are involved in the creation of national version of INSPIRE, the various implementing regulations will be gradually implemented into national legislation.

LITERATURE

[1] Ing. Tuchyňa, Ing. Kazárová, Ing. Cimerman., " Enviromagazín 5/2004".
[2] Mgr. Michal Štefanek., (2006, august). „EnviroInfo“, Používateľská príručka v 1.0

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Large-scale Research Infrastructures to assist in understanding and managing our Planet's Biodiversity

Wouter Los[a]

[a] *Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Science Park 404, 1098 XH Amsterdam, The Netherlands (w.los@uva.nl)*

Abstract: Understanding the biodiversity system requires a "systems biology" approach in addition to the reductionist method with experiments on small components of the system. This methodological approach requires the availability of large-scale databases based on intensive observations and measurements, together with advanced analytical and modelling software in a high performance computing environment. The European Strategy Forum on Research Infrastructures selected the 'LifeWatch e-science and technology infrastructure for biodiversity research' as a promising development to construct facilities for this new approach. It will allow collaborative networks the integrated use of public resources, such as data repositories, sensors and computational capacity through a service-oriented architecture. Likewise the biodiversity policy domain will benefit from new dedicated decision-support services.

Keywords: Research infrastructure; Biodiversity observatories; Environmental sensors; eScience

1. INTRODUCTION

Understanding and managing changes in biodiversity is crucial for the sustainable development and exploitation of Europe's resources in relation to climate change and global change in general. Some of the impacts of climate change such as changing precipitation patterns, droughts and fires, frequency and intensity of storms, sea level rise, acidification of the oceans and so on will inevitably lead to changing species distributions and migration of whole biogeographical zones. The rate of such changes is known already for a number of well-studied groups such as birds and butterflies or forests and agricultural crops but not for many others, including most marine populations, many pests and parasites and invading species. Changes in biodiversity are having serious effects on the capability of ecosystems to provide essential services. This in turn will affect the quality of life of citizens and social and economic aspects of sustainable development.

We face however serious problems in the understanding and managing of biodiversity. The biodiversity system is complex and cannot be described by the simple sum of its components and relations. The functioning of this system, its response to pressures, and feed-back mechanisms to sustain stability in the system are hard to unravel in experiments. The data, information and knowledge required to address these problems are scattered and incomplete. In particular, the long-term data required to understand the response of ecological systems to social, economic and global changes are rarely available and the lack of spatial comparability of data across the globe severely inhibit our capacity to develop cost-effective policy responses. Without another approach it would also not be possible to provide scientific evidence to the policy domain for making management decisions.

2. NEW METHODOLOGICAL APPROACHES

2. 1 The reductionist method fails in studying whole biological systems

The science of biology and also of biodiversity has made progress with reductionist approaches by studying isolated separate components of the biological system. Much knowledge has been generated by manipulating a few parameters in experimental set-ups. This reductionist approach fails however in capturing a wider picture of the full complexity of biological systems. This holds for all levels of biological life, on the cellular level, the organism level, and the ecological level. Especially the biodiversity system is very complex since its covers subsystems on all the levels from genes to species up to ecosystems. In addition, these subsystems operate on different spatial and temporal scales which can not easily be interrelated. It makes no sense to simply extrapolate from - for example - soil microbiological processes on the cubic centimetre scale up to processes on the scale of river basins, and vice versa. The same holds for processes on the ecological time scale (from a few hours to hundreds of years), compared to the evolutionary time scale (millions of years). Biologists have designed alternative methodological approaches to understand biological processes at the level of whole systems.

2.2 Systems biology

Biological systems are characterized by self-organization resulting in a high variety of diversity and complexity in order to adapt to external constraints (environments). This is known for the variation of configurations of single proteins up to ecological communities [Guill and Drossel, 2008]. It is very often not possible to understand a biological system by extrapolating from the known behaviour of single units which constitute the system. It is in such situations inevitable to apply different methodological views on the system by also analysing the correlation properties of all (ensembles of) units with statistical methods [Dhar, 2007]. Systems biology is based on such methods with intensive computation. High throughput technologies assist in detecting patterns of strong correlations with evidence for "collective organisations", which in turn can be further analysed to retrieve the processes resulting in such patterns [Conti et al, 2008].

This methodological approach assumes the availability of large-scale databases with observations and measurements of the components of biodiversity, together with advanced analytical and modelling software. In addition it requires computational capacity in order to run the demanding statistical work flows on the huge data sets. These requirements together define an infrastructure environment which brings together data, software and computation facilities at an appropriate integrated large scale. Scientists have worked the last two decades to develop the components of such an infrastructure for biodiversity research [Los and Hof, 2007]. Expected further technological progress in the near future opens the opportunity to integrate the components and scale these up to a service-oriented operational research infrastructure to support biodiversity science and policies.

3. THE EUROPEAN STRATEGY ON RESEARCH INFRASTRUCTURES

3.1 European Roadmap for Research Infrastructures

In order to explore and promote the building or upgrading of research infrastructures of European significance, the EU countries established the European Strategy Forum on Research Infrastructures (ESFRI). The preparation of the ESFRI Roadmap with promising new European research infrastructures was a main activity of the Forum [ESFRI, 2006]. The European Strategy Forum on Research Infrastructures selected a few large-scale research infrastructures as essential new facilities to promote breakthroughs in science and to support wise policy decision-making. Some of these are supported to prepare for the construction of the infrastructure and finally get enough countries together willing to invest in the new infrastructure. Other initiatives are selected as emerging infrastructures, and are in the stage to design the architecture of the planned facilities. The positions of a few initiatives related to biodiversity research are shown in Figure 1.

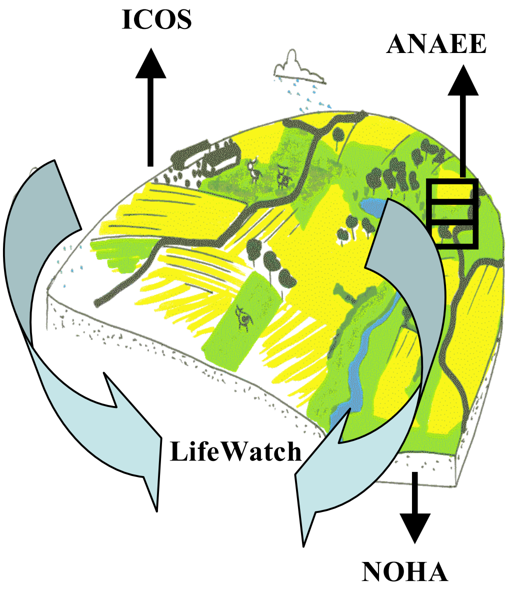


Figure 1

Preparing:

- *LifeWatch*, e-Science and Technology Infrastructure for Biodiversity Research;
- *ICOS*, Integrated Carbon Observatory System;

Designing:

- *ANAEE*, European Research Infrastructure for the Analysis and Experimentation on Ecosystems;
- *NOHA*, Network of Hydrological Observatories for Water Resources Research in Europe.

Each of these emerging research infrastructures has a complementary focus on the natural system, but they interact to achieve the integration of the infrastructures at landscape and regional scales. This applies for their operations with respect to data capture and transmission, data management and data integration, and integrated modelling.

3.2 LifeWatch infrastructure for biodiversity research

LifeWatch is an e-Science research infrastructure designed to explore, describe and understand the complexity of biodiversity [LifeWatch, 2008]. It supports research on biodiversity and on the environmental and anthropogenic impacts affecting biodiversity. As such it is relevant for human society in the areas of human health, food, bio-energy and the use and conservation of nature. The LifeWatch infrastructure will allow scientific research teams to create '*e-Laboratories*' or to compose '*e-Services*'. All public resources, such as data repositories, computational capacity and sensors are available through a service-oriented architecture. The standards-based architecture also allows for linkages to external resources and associated infrastructures. As such, LifeWatch is as an example of a new generation of research infrastructures existing in a cooperating fabric of such infrastructures.

One of the external resources on which LifeWatch will build its services, is the Global Biodiversity Information Facility (GBIF). Through web-services, users may benefit from GBIF's data while doing their scientific modelling in the LifeWatch e-Laboratory environment. The same holds for other resources, such as distributed computational capacity. As a distributed research infrastructure LifeWatch also benefits from the data generated in many research networks, including the EU funded Networks of Excellence in biodiversity research. Users may share their data with other users. For species-level data LifeWatch will direct the data automatically to the data sharing functionality of GBIF.

As an infrastructure connecting or integrating the capabilities of other facilities, LifeWatch itself will contribute to developing GEO BON, the biodiversity and ecosystem observation component of GEOSS, the Global Earth Observation System of Systems [Scholes et al, 2008].

The architecture of LifeWatch is modular with several modules connecting to external resources, supporting Grid e-infrastructures and user communities (Figure 2).

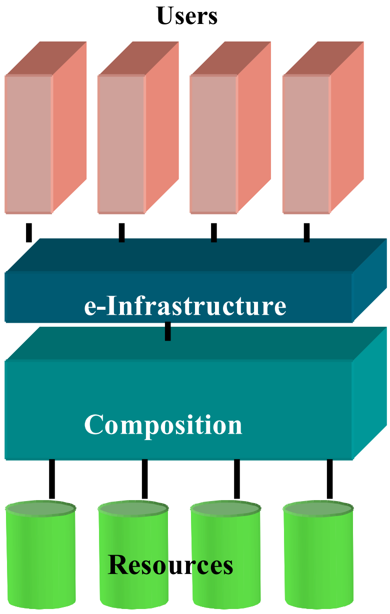


Figure 2

- The *User Layer* provides domain-specific presentation environments for the control and monitoring of tasks.
- The *Composition Layer* supports the intelligent selection and combination of services in order to complete tasks.
- The e-*Infrastructure Layer* provides mechanisms for enabling the sharing of the resources as generic services in a distributed environment spread across multiple administrative domains.
- The *Resource Layer* contains a variety of conditional resources, such as data repositories, sensor and human observation networks, modelling software and computational capacity.

4. INNOVATION TO SUPPORT BIODIVERSITY RESEARCH AND POLICIES

4.1 Biodiversity data and environmental sensor networks

Infrastructure services for scientists and the policy domain depend on the availability of biodiversity data, and such data should be accessible in open and user-friendly formats [Arzberger, 2004; ENBI-GBIF statement, 2005]. Such data are of various kinds and basic data include for example: species occurrences (in time and space), their genetic variation, bio-mass, and carbon and nitrogen fluxes. Such data should not only originate from a few locations, but preferably represent all parts of a fine-mazed grid covering a continental area or sea, or even better the whole planet. In this respect, Europe is in the relative good position with a variety of networked organisations which cooperate on data capture with common protocols. Examples are the Long-Term Ecological Research Network Europe (for terrestrial monitoring) and its associated Network of Excellence AlterNet, the Marine Research Stations Network and its MARBEF network, and the Consortium for European Taxonomic Facilities and its European Distributed Institute for Taxonomy. A number of EU Networks of Excellence in biodiversity research is integrating the data generating activities to a level that these can operate as a crucial data resources in the LifeWatch research infrastructure. Other networks bring together the observation activities of ten thousands of "citizen scientists"; experienced volunteers who send their (GPS supported) species observations in defined areas at a regularly basis to a central shared data repository [Irwin, 2001].

All such monitoring activities benefit from new technological solutions to collect much more data of higher quality and at lower cost. Autonomous operating wireless environmental sensors or smart hand-held devices to transmit observations will accelerate data generation with more reliability and at the long term also cheaper. With such equipment it becomes possible to design ubiquitous sensor networks, or interacting human observation networks. Such intelligent networks are capable to notice observation anomalies, to automatically increase the observation frequencies, to act as an early-warning system and thus provide much more value from the network. A condition is the availability of dense antenna networks to receive and transmit the data with capabilities for autonomously generating spatial correlations between the data [Nieberg, 2003]. The mobile telephone networks are dense but are missing autonomous intelligent capabilities. This is different for new research antenna networks as these are implemented or designed for radio astronomy research. An example is the new radio telescope of the LOFAR (Low Frequency Array) project with tens of thousands of antennas that are connected to each other with a large ICT infrastructure. LOFAR-Agro will make use of this infrastructure and has chosen as its first application the measurement of the micro-climate in potato crops. This information will be used to improve the advice on how to combat phytophthora within a crop, based on the circumstances within each individual field [Lokhorst, 2008]. The international Square Kilometre Array project can likewise provide the antenna backbone for environmental sensors [SKA, 2009]. The LifeWatch infrastructure project is exploring these opportunities.

4.2 Collaborative networks

Organizations involved in biodiversity, e.g. the biodiversity data providers, laboratories, universities, conservation groups, etc. are increasingly involved in collaborative activities with other organisations in the biodiversity domain, as well as with organizations that do not belong to the biodiversity domain. Typically, such collaboration aims to perform some monitoring, research, or development task at the level of a specific geographical region, a country, or the entire continent, where they need to share their competencies and resources and cannot perform it alone. The LifeWatch preparatory project also designs ICT supported mechanisms for goal-oriented temporary collaborations distributed around Europe and the World. [Camarinha-Matos and Afsarmanesh, 2008]. Such Temporary Collaborative Networks may consist of a network of legally independent organizations that come together and share skills, abilities, and resources to achieve common goals, from jointly preparing a proposal or a bid, to jointly perform the tasks needed to satisfy an opportunity/aim in the market/society or to innovate, while using the computer networks as the base communication/interaction infrastructure (Camarinha-Matos and Afsarmanesh, 2004).

A wide variety of issues must be resolved, considering especially the huge number of organizations and the complicated issues must be addressed to support their collaboration (interactions, roles, trust), as well how they collectively can work together by accessing and sharing data, software and computation. The traditional way of engineering infrastructures, relying on dedicating the hardware for a single purpose and a single user group, must be replaced by sharing resources in virtual environments. With virtualization the different operating systems of collaborators may share the resources in a secure and efficient way. It is so possible to establish various dedicated environments for different applications in biodiversity science and for policy services. Virtualization offers uniform interfaces and allowing for easily modifications and upgrades, even without noticing by the user [Baldridge, 2003].

5. CONCLUSIONS

In the last two decades researchers have developed and improved observation techniques, database interoperability and analysis algorithms to support biodiversity research. This opens the way to new methods to analyse and model the biodiversity system, in addition to controlled scientific experiments on restricted parts of the system. Such a "Systems Biology" approach requires the integration and upscaling of the separate components of data generation, data accessibility, software, and the intelligent combination of these services. Now the time is right to apply the emerging e-infrastructure architectures to facilities for biodiversity research. We see the rise of an integrated grid of sensors, observatories, data repositories, software tools, high performance computers, and collaborative networks in virtual environments.

This will dramatically change research and will lead to a mind-shift of users and encourage new young scientists to exploit the new opportunities. Likewise the biodiversity policy domain will benefit from new dedicated decision-support services. The LifeWatch e-science and technology infrastructure for biodiversity research is preparing the actual construction of such facilities.

ACKNOWLEDGEMENTS

The LifeWatch Preparatory Project is supported by the European Commission with FP7 Grant Agreement No 211372-LIFE WATCH.

REFERENCES

Arzberger P., and P. Schroeder, A. Beaulieu, G. Bowker, K. Casey, L. Laaksonen, D. Moorman, P. Uhlir, P. Wouters. Promoting Access to Public Research Data for Scientific, Economic, and Social Development. *Data Science Journal* 3, 135-152 (2004).

Baldridge, K. and P.E. Bourne. The New biology and the Grid. *Grid Computing*, pp 907-922. Wiley Series in Communication Networking & Distributed Systems. On-line ISBN 9780470867167 (2003).

Camarinha-Matos, L.M. and H. Afsarmanesh. The emerging discipline of collaborative networks. *Proceedings of PRO-VE'04 - Virtual Enterprises and Collaborative Networks*, Kluwer Academic Publishers, ISBN 1-4020-8138-3, pp 3-16, (2004).

Conti, F., and M.C.Valerio, J.P. Zbilut, A. Giuliani, Will systems biology offer new holistic paradigms to life sciences? *Syst Synth Biol* 1(4): 161-165 (2007).

Dahr, P.K., The next step in biology: a periodic table. *J Biosci* 32:1005-1008 (2007).

ENBI – GBIF Digitisation Statement on Free and Open Data Access (2005), http://circa.gbif.net/Public/irc/gbif/pr/library?1=/access_statement/_EN_1.0_&a=i.

ESFRI. European Roadmap for research Infrastructures – Report 2006. Office for Official Publications of the European Communities, ISBN 92-79-02694-1, http:cordis.europa.eu/esfri/ (2006).

Guill, C, and B. Drossel, Emergence of complexity in evolving niche-model food webs. *J Theor Biol* 251: 108-120 (2008).

Irwin, A. Constructing the Scientific Citizen: Science & Democracy in the Biosciences, *Public Understanding of Science* vol.10, pp.1-18 (2001).

LifeWatch, e-science and technology infrastructure, http://www.lifewatch.eu (2008).

Lokhorst, C. and Eijk, O.N.M. van; Frumento, E; Glisenti, Fulvio; Goense, D.; Hogewerf, P.H.; Ipema, A.H.; Mazzu, Marco; Mol, R.M. de; Savio, Walther; Wells, Hannah. Wireless Access to Sensor Populations: Potential applications for herd health control , eldery care and road transport. *Proceedings 3rd ECPLF conference 3-6 june 2007, Skiathos, Greece* (2008) .

Los, W and C.H.J. Hof. The European Network for Biodiversity Information. *Biodiversity databases*. pp. 5-12. Editors G.B. Curry & C.J. Humphries. Taylor & Francis / CRC Press (2007).

Nieberg, T and S. Dulman, P. Havinga, L van Hoesel, Jian WU. Collaborative algorithmes for communication in wireless sensor networks. *Ambient intelligence: impact on embedded system design*. pp. 271-294. Kluwer Academic Publishers, ISBN:1-4020-7668-1 (2003).

Scholes, R.J., and G.M. Mace, W. Turner, G.N. Geller, N. Jürgens, A. Larigauderie, D. Muchoney, B.A. Walther, H.A. Mooney. Toward a Global Biodiversity Observing System. *Science*, 321: 1044-1045 (2008).

SKA. Square Kilometre Array project, http://skatelescope.org (2009).

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Electronic Data Management

F. Mochty[a]
[a] *Federal Ministry of Agriculture, Forestry, Environment and Water Management, Vienna, Austria (franz.mochty@lebensministerium.at)*

Abstract: ***E**lectronic **D**ata **M**anagement* is a project of the Federal Ministry of Agriculture, Forestry, Environment and Water Management within the Austrian eGovernment initiative. The main objective of EDM is to support public administration and to reduce costs especially for recording and reporting obligations for companies by modern eGovernment techniques. In the future also applications and approval processes for permits in the waste management sectors will be facilitated. EDM is based on a central master data registry and provides a single-sign-on access to several specific applications dealing with recording and reporting of environment-related data according to the existing legal framework. Special emphasis is placed on the use of international cross-sectoral standards for identification systems and data models in order to provide interoperability and future safety. Thus the reuse of data will be possible, multiple data collection and processing can be avoided and administrative processes will be facilitated. By data evaluation across the different reporting obligations a better overview on the condition of the environment can be achieved and the effectiveness of legal measures can be estimated.

***Keywords**:* Environmental data; eGovernment, recording and reporting obligations, central master data registry, waste management, edm.gv.at.

1. BACKGROUND

Business and public administration are confronted with numerous reporting obligations related to environmental issues. Just to name some examples:

- EC Waste Framework Directive
- EC Waste Shipment Regulation
- EC Directive on Hazardous Waste
- EC Landfill Directive and Decision
- EC Waste Incineration Directive
- EC Directive on Waste Electrical and Electronic Equipment
- EC Directive on End-of-Life Vehicles
- EC Battery Directive
- EC Regulation on the European Pollutant Release and Transfer Register
- EC Water Framework Directive
- EC Directive on the Scheme for Greenhouse Gas Emission Allowance Trading
- Implementation pursuant to the Kyoto Protocol 1997 (HFC Ordinance)
- ……….

This – non-exhaustive – list gives an impression of the resulting administrative expenses. Basically every report consists of two kinds of data: data describing who is reporting about what (master data) and data specifying the relevant matter, so-called variable data. The master data comprise basic properties which don´t change regularly. These data are the basis of almost every kind of reports and administrative processes. On the contrary, variable data describe changing properties with a specific time reference (e.g. tons of CO_2 emitted in a specific year, tons of waste treated last year, the exceeding of emission limit values in 2008, etc.) and thus have to be adapted for each specific report.

EDM aims at the simplification and streamlining of administrative needs and costs without reducing standards and quality in environmental law enforcement. Therefore drawing up a central registry for master data providing its content for as many reports as possible is a main objective of EDM. For the reuse of environmental data cross-sectoral harmonization of the content and structure of recording and reporting obligations is essential. By these measures the expenses for maintenance of the master data for the different reports and administrative processes can be reduced, recorded data can be reused to the extent possible and multiple data collection can be avoided. To overcome this challenge the EDM Program is based on the following principles:

- Implementation of real EDI (Electronic Data Interchange)
- Electronic interchange of structured data (quantified or classified) using approved message standards directly between IT applications with a minimum of human intervention
- One-time manual input of data into the electronic system – where the information is initially collected – then only electronic transfer and processing
- To the extent possible use of already available data (eg waste management data derived from procurement and accounting)
- Integration of the entire business processes into EDI adjusted to the economic and technical capacities of the participants

Putting these principles into practice leads to the following consequences:

- Unique data type definitions for all applications
- Cross-sector data and message definitions
- Utilisation of internationally unique, cross-sector identification systems (the GS1-ID-scheme is used for all identifiable objects within EDM)
- International standardisation (UN/CEFACT) of data models and message structures (the EDM data model used for waste movement is due for approval by UN/CEFACT this year)
- International co-operation (e.g. transfrontier shipment of waste)

To earn the best returns the principles given above have to be considered already during drafting legislation on a national and international level. European legislation should also follow these principles, especially the recognition of international cross-sectoral standards is essential.

SISE (Information Space in Europe for the Environment), SEIS (Shared Environmental Information System), INSPIRE (Infrastructure for Spatial Information in the European Community) are important examples where the application of these principles is a key issue.

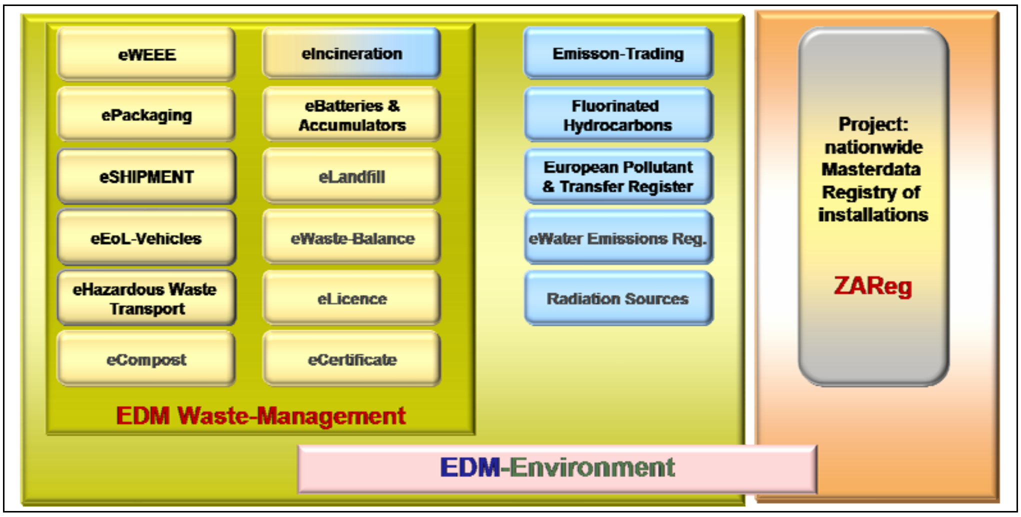


Figure 1: The Austrian EDM Program

2. REGISTRY OF MASTER DATA

EDM comes up with a central registry of master data using the GS1-IDs (Global Location Numbers) for persons, locations/facilities and installations and in this way forms the basis for periodic reports stemming from various legal acts in the environmental field. Using the already widely familiar GS1-standard among enterprises resource planning should guarantee a maximum of interoperability and acceptance.
For the time being the user authentication is realized via user-name and password. Registration is mandatory for individuals and legal entities (like companies) as far as they are obligated parties according to different legal regulations (e.g. Waste Management Act, Emission Trading Act). At the current state approx. 30,000 persons (individuals and legal entities) are registered. At a final stage 300,000 registered persons are expected. To prevent data redundancy and moreover data inconsistency the master data registry enables synchronisation with existing Austrian eGovernment Registries (e.g. the Company Register).
Data are provided by registered persons and scrutinized and completed by competent authorities via a web application. The covered data consist of details such as name, type of enterprise, industrial sector, address, addresses of service, contact persons, related reporting and recording obligations as well as spatial information such as plot numbers and polygons of single installations generated via a web-GIS. Public authorities feed in data of company-specific permits and installation-related permits. Public queries are available for example to find out which company is authorized to collect or treat a certain hazardous waste in a specific region. In a forthcoming version companies will also be enabled to apply for various permits via EDM.
The master data registry is designated to be used within public administration as well as for business related purposes especially in the waste management sector (e.g. logistics, procurement). Due to its cross-linked structure data security is an important issue.

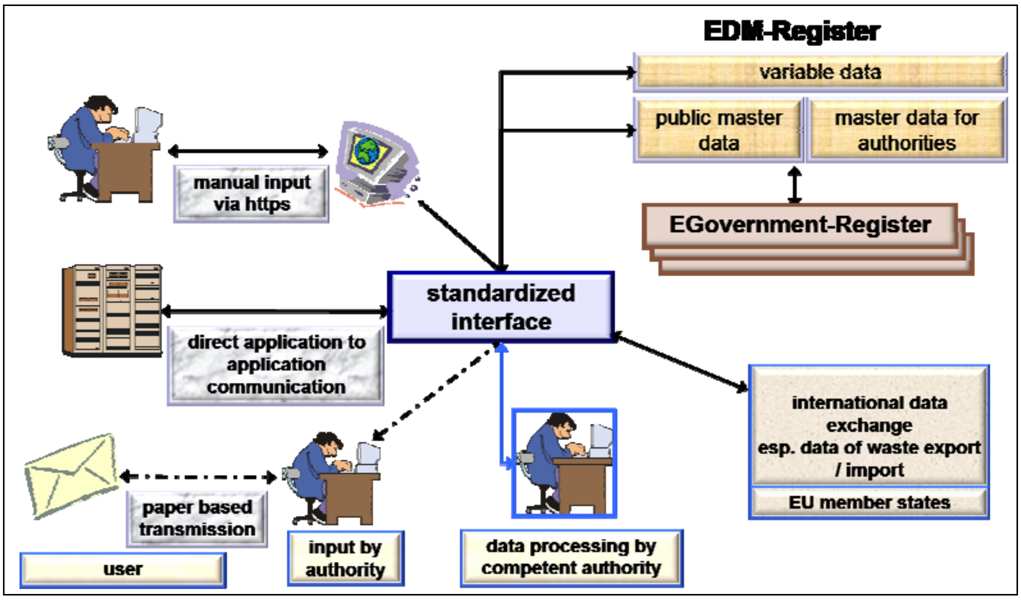


Figure 2: Information Flow with National EDI System “EDM”

3. REPORTING APPLICATIONS

3.1 Examples of Applications online

EDM Portal: edm.gv.at is a gateway for public use as well as the single point of access for persons registered and public authorities. Basic information about EDM and general master data are available with the help of a public query and (to a certain extent) downloadable.

The public sector of the *EDM Portal* also allocates descriptions of service interfaces and tables of reference data used for electronic data interchange by different applications hosted within the EDM service framework.

eHazardous Waste Transport: Council Directive 91/689/EEC on hazardous waste and the associated national legislation in Austria requires the tracking of every single transport of hazardous waste by consignment notes. Data concerning the type and mass of waste, IDs of sender and consignee, location of hand-over and disposition, and date of transport must be reported to the competent authority within 3 weeks after the transport took place. Approximately 500,000 messages per year are processed via upload of XML files and data are held available for 9 provincial authorities in Austria. One third of these messages are already submitted through direct electronic data interchange (EDI) between ERP systems of waste management companies and EDM, resulting in noticeable cost reduction. For the coming years *eHazardous Waste Transport* focuses on strengthening EDI and developing an integrated electronic recording/reporting process (paperless).

eWaste Shipment: EC Regulation 1013/2006 establishes procedures and control regimes for the shipment of waste between Member States as well as import, export to and transit through Member States depending on the origin, destination and route of the shipment, the type of waste shipped and the type of treatment to be applied to the waste at its destination. To simplify this rather time consuming, paper and fax based process the European project EUDIN for electronic processing of data between Member States has been started with The Netherlands, Belgium, Germany, and Austria. Currently the project results are laid down in a technical paper describing the interface between national applications and the so-called "message broker" and the underlying data model will be adopted by UN/CEFACT this year. The EUDIN Project team is looking forward to boost its activities during the near future. In the meantime Austria has successfully implemented a bilateral (Baden-Württemberg, Bavaria, Austria) pilot application handling the relevant data of more than 90,000 transfrontier transports of excavated material stemming from a road tunnel construction site. A rough estimate promises EUR 0.5 mio cost savings each for public authorities and for the companies involved.

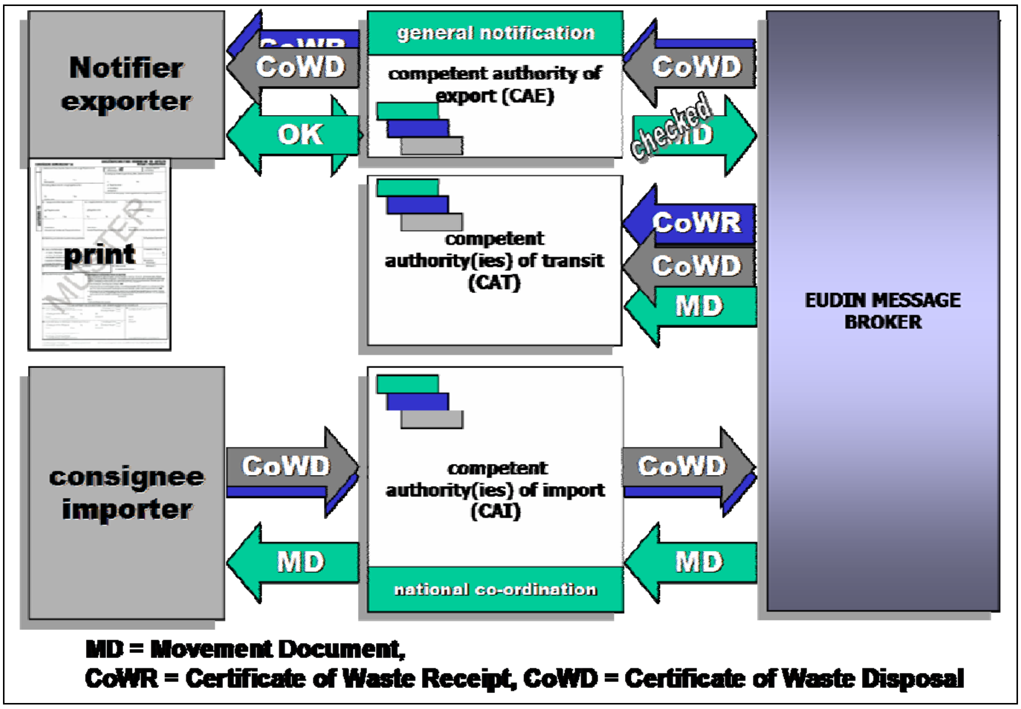


Figure 3: European Data Interchange for Waste Notification Systems

ePRTR: According to the EC-Regulation 166/2006 concerning the establishment of a European Pollutant Release and Transfer Register specific information on releases of pollutants to air, water and land and off-site transfers of waste and of pollutants in waste water have to be reported by operators of facilities carrying out specific activities [EC, 2006]. *ePRTR* is a web-application for obligated parties and authorities. Data already submitted to one of the EDM applications (due to another legal context) are used for the ePRTR report (e.g. mass and type of waste handed over) in order to avoid data redundancy. The application supports validity checks carried out by competent authorities.
Other applications already in use are eWEEE and eBatteries, offering all kinds of reporting activities in the context of used electric and electronic equipment, batteries and accumulators acc. to EC Directives 96/2002 and 12/2008. 95% of the messages are handled via webservice. Moreover, EDM provides reporting applications for data about emission trading, emissions of waste incineration plants and thermal power stations, release of fluorinated hydrocarbons, treatment of end-of-life-vehicles, treatment of packaging material, and waste delivered to landfills.

3.2 Applications in preparation

Within the next years EDM will broaden its services to enable reports on emissions to surface water. Also a registry of radiation sources will be integrated into the EDM master data registry. In the field of waste management a new application will be launched in 2010 by means of which waste treatment facilities will report their input/output balances (eWaste-Balance).
Implementing the Council Decision of 19 December 2002 (EC 2003/33) establishing criteria and procedures for the acceptance of waste at landfills (acc. to Article 16 of and Annex II to Directive 1999/31/EC) the Austrian Government issued a new Austrian landfill ordinance. A special focus is laid on the rather complex requirements concerning waste characterization and criteria and procedures for the acceptance of waste at landfills. Right from the design of the regulation an electronic support system was developed in parallel to facilitate the enforcement of the regulation. An electronic expert system will support the development of sampling plans according to the specifications of the ordinance, the evaluation of the data of the basic characterization, the check if the results comply with the limit values and acceptance criteria of a specific landfill, the on-site verification at the landfill, and the inspection by the competent authority.

4. FUTURE PROSPECTS

Beside the establishing of an electronic expert system for waste characterization and waste acceptance at landfills the integration of existing master data in the Austrian provincial governments (according to industrial and water legislation) and the cooperation with these provincial registers is a serious challenge within the next years. The integration of other Austrian eGovernment register like the building and household register for address data will further improve the data quality. Furthermore, EDM will facilitate the application for a permit to collect and treat waste and support the governmental approval process.

5. CONCLUSIONS AND RECOMMENDATIONS

There is a high potential of cost savings by establishing a comprehensive eGovernment system for all environment-related data and administration processes. Cost reduction at public administration and the business sector, the reuse of data, the avoidance of multiple data collection and processing, the facilitation of administrative processes, the support of difficult decision processes by electronic expert systems as well as a better overview on the condition of the environment and an estimation of the effectiveness of legal measures can be achieved by an integrated electronic register in the environmental sector.
Prerequisites are cross-sectoral data harmonisation on an international basis, system-wide unique identifier, internationally standardised data and message structure, close cooperation between public administration and business sector, and harmonized reporting obligations.
To secure the potential savings support by the relevant bodies of the EU is essential.

REFERENCES

European Commission, Guidance Document for the implementation of the European PRTR, *http://eper.eea.europa.eu/eper/documents/GuidanceDocs/*, 2006.

European Commission Enterprise and Industry Directorate-General: *Streamlining and Simplification of Environment-Related Regulatory Requirements for Companies – Final Report of the Best Project Expert Group*, May 2006 http//:europa.eu.int/comm/enterprise/enterprise_policy/best/best_procedure.htm

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Integrating Environmental Research Data into the Shared Environmental Information System

P. Mooney[a], **P. Corcoran**[b] ,**A.C. Winstanley**[b]

[a] *Environmental Research Centre, Environmental Protection Agency, Richview, Clonskeagh, Dublin 14. Ireland (p.mooney@epa.ie)*

[b] *Department of Computer Science, National University of Ireland Maynooth, Co. Kildare. Ireland (padraigc@cs.nuim.ie, adam.winstanley@nuim.ie)*

Abstract: The paper describes efforts underway at the Irish Environmental Protection Agency (EPA) to ensure that environmental research data generated by the EPA's research funding programmes are properly managed with a view of future integration of these data streams into SEIS. The paper argues that the flows of data and information from environmental research will greatly enhance the usage and application of SEIS by assisting European citizens, stakeholders, decision makers, and policy makers obtain access to research data. The thematic, geospatial, and temporal heterogeneity of data and information flows from environmental research will provide SEIS with a rich geoscientifc resource.

Keywords: Environmental research data; Environmental information; Shared Environmental Information System;

1. INTRODUCTION

Scientific research is generating vast and ever increasing quantities of information. In the age of the Internet the provision of free and efficient access to environmental information, including scientific publications and original data, will be the key for sustained progress on environmental topics [ERC, 2007]. Universities and public research institutions perform more than 35% of all research undertaken in Europe [EC, 2007]. Research directly connected with environmental assessment and sustainability is a significant part of this output. Access to the outputs of this research is often difficult to obtain. Intellectual property rights (IPR) issues, capacity and resource problems experienced by the researchers technical issues such as interoperability, and cultural differences to sharing and exchanging data, all combine to compound the problem. The European University Association (EUA) responded to the green paper on “The European Research Area: New Perspectives” by stating that the full potential of ICT technologies can be used more effectively to enhance “open access” to scientific results arising from publicly funded research to a wide variety of stakeholders including the public and the research communities. Measures are steadily being introduced to address these issues. For example in August 2008, the European Commission launched an open access pilot scheme for FP7. Under this pilot scheme, grant recipients in seven designated areas will be required to: deposit peer-reviewed research articles or final manuscripts from FP7 projects into an online repository and to make their best efforts to ensure open access to these articles within either six or twelve months after date of publication. Unfortunately this does not include access to the raw data generated by the projects or derived data products. Access to raw data and information related to these publications is needed not only for independent verification of results but also for secure longer-term preservation and "fresh analysis and utilization of the data" [ERC, 2007].

In February 2008 the European Commission issued a Communication [EC, 2008] entitled "Towards a Shared Environmental Information System (SEIS)". The purposes of SEIS is to improve the efficiency and streamline the European systems for collecting, analysing, exchanging, and reporting environmental information and data for the design and implementation of environmental policy. An underlying aim of SEIS is to move away from paper-based reporting to a ICT system where information is managed as close as possible to its source as possible and this information is subsequently made available to users in an open and transparent way. While SEIS will predominantly work with data and information products generated by Member States for National and EU level reporting obligations this paper outlines the important advantages of integrating environmental research data products into SEIS.

EPA STRIVE is a €100 million programme to fund environmental research and innovation in Ireland. STRIVE runs from 2007 until 2013. Many different types of projects will be funded ranging from short desk studies to large multi-annual capability development projects involving several organizations. STRIVE is likely to involve over 1,000 researchers and industrial innovators and be highly relevant to the key national and international environmental priorities in Ireland. In following its mission to become a key source of environmental data and information for Ireland the EPA promotes the sharing of the outputs of STRIVE research to ensure that these can be utilized to the maximum possible extent by Irish stakeholders, the international research community, other EU member states, interested parties, and the wide general public. Research data represents one of the fundamental outputs of scientifc research. In various contractual agreements with funded researchers the EPA makes their expectation very clear that researchers should make their data widely available as soon as feasible and by placing as few restrictions upon this access as possible.

1.1 Current Status

Currently at the EPA the outputs from the STRIVE programme are managed in such a way that they can be woven seamlessly into the fabric of SEIS in the future. In Ireland environmental research funded by the EPA has played a pivotal role in assisting Irish efforts to meet EU reporting obligations and legislation. Some prominent examples include research projects which:

- Supported the Irish role in the EU Intercalibration exercise of Biological Quality Elements (BQE) for the assessment of water quality in transitional and coastal waters in Europe under the Water Framework Directive;
- Supplied scientific data to underpin appropriate measures or actions that might be used in the implementation of national policy for reducing phosphorous (P) and nitrogen (N) losses to water from agricultural sources;
- Examined the European Regional Air Pollution and Simulation model (RAINS) whose outputs are used in the setting of emissions ceiling targets for 2020 under the EU National Emission Ceilings Directive.

It is important the outputs such as those above *and* the raw data and information that corresponding to them are available to a wide range of stakeholders and policy makers through SEIS. It is also important that they are available for planning and policy making at a regional and local level or for Strategic Environmental Assessment.

To better coordinate the management of research data and information outputs from STRIVE the Secure Archive For Environmental Research Data (SAFER-Data) [SAFER, 2008] has been developed. SAFER-Data is a web-based system providing access to the outputs from the STRIVE programme and from previous EPA funding programmes. The

SAFER system performs a dual role. On one hand it provides the '*pull*' tools for researchers to:

- Securely archive their raw data and information for long-term preservation;
- Document these resources with INSPIRE compatible ISO19115 metadata;
- Manage and update these resources going forward;
- Specify the levels of public access to their resources.

On the other hand it provides the '*push*' tools for the EPA to disseminate these data and information resources to as wide a stakeholder and user audience as possible in a flexible and accessible manner. An in-house EPA team of scientific officers and specially chosen scientific experts review the suitability of data resources and derived information products before they are made publicly available. The data providers and data owners are pro-actively engaged in this process to ensure that SEIS Principle 6 "Information should be fully available to the general public, after due consideration of the appropriate level of aggregation and subject to appropriate confidentiality constraints, and at national level in the relevant national language(s)" [EC, 2008] is successfully implemented. This principle is particularly important in the case of some Biodiversity projects and projects with a social science aspect where the privacy of endangered species or individuals must be protected. Another key advantage of our approach is that SAFER is not restricted to geospatial data. The EPA look to make non-spatial and non-quantitative environmental research data and information available for the benefit of those who require a range of data types such as Strategic Environmental Assessment.

2. CONCLUSIONS

At this early stage in their evolution it is difficult to visualize the precise "look-and-feel" of the SEIS or SISE (Shared Information Space for the Environment). However given SEIS principles (1) and (2) it is clear that the structural design of SEIS will involve connecting national and regional data providers and data sources through international standards for data exchange and web-services such as Web Map Services or Web Catalogue Services. SAFER-Data shall implement interfaces such as: Open Archives Initiative for Metadata Harvesting; GeoRSS and RSS Feeds; Web Map Services where appropriate; Catalogue Web Services. Some of the international standards for data exchange will involve conversion of certain data resources to KML allowing users to visualize data on platforms such as Google Earth. It will be important to ensure that there are clear and effective procedures in place to determine which data flows and outcomes from the research programme are of the highest value. It is clear that not all of the data generated will have national or EU-level reporting significance. However the actions on environmental research and capacity building outlined in STRIVE [EPA, 2007] will mean that a significant and growing portion of this data will help to inform policy makers and provide information for sustainable environmental management. The authors believe that preparing data and information from the environmental science research community for inclusion in a SEIS would be a significant step towards "stimulating the development of a 'continuum' of accessible and interlinked scientific information from raw data to publications" [EC, 2007] while also providing SEIS with access to a rich geosceintific resource. SEIS can be "brought alive" by demonstrating how environmental information and data are used by policy makers and the public. SEIS can add value to environmental research by providing a European wide infrastructure for access and dissemination of research outputs. In return the environmental research community can add value to SEIS by contributing new software tools and methodologies for access to research data while utilising to access other data flows.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the financial support for this work from the Science, Technology, Research and Innovation for the Environment (STRIVE) Programme 2007–

2013. The programme is financed by the Irish Government under the National Development Plan 2007–2013. Dr Peter Mooney is the Principal Investigator on this project which is titled “Geoinformatics Services for Improved Access to Environmental Data and Information”.

REFERENCES

ERC., European Research Council: ERC Scientific Council Guidelines for Open Access, *http://erc.europa.eu/pdf/ScC_Guidelines_Open_Access_revised_Dec07_FINAL.pdf* , 2007.

EC, Green Paper The European Research Area: New Perspectives. Commission of the European Communities COM(2007) 161 Final. Brussels, April 2007.

EC, Towards a Shared Environmental Information System (SEIS). Commission of the European Communities COM(2008) 46 Final. Brussels, February 2008.

EPA, STRIVE – Science Technology, Research and Innovation for the Environment 2007 – 2013. The Environmental Protection Agency Ireland. PO Box 3000, Johnstown Castle Estate, Co. Wexford Ireland. Available online at *http://www.epa.ie/downloads/pubs/research/* (Last Checked Sept 2008) Oct 2007.

SAFER, The Secure Archive For Environmental Research Data – SAFER-Data *http://erc.epa.ie/safer* (Last Checked Sept 2008) September 2008

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Access to Water Data in Austria - Water Information System Austria

Harald Marent[a] **and Stephan Nemetz**[b]
[a] *Federal Ministry of Agriculture, Forestry, Environment and Water Management, Austria (Harald.MARENT@lebensministerium.at)*
[b] *Umweltbundesamt, Austria (stephan.nemetz@umweltbundesamt.at)*

Abstract: In 2003 the establishment of the Water Information System Austria – WISA was launched, its primary goal being to ensure the successful implementation of the Austrian Water Act (Wasserrechtsgesetz, WRG 1959) as applicable, but also the requirements of the Austrian Environmental Information, the access to environment data has become more important than it was ever before. WISA has to fulfil the needs and requirements of a modern management tool in water management on an interregional level. Thereof it depends how the data are accessible or presented and the aggregation of data shall correspond with the user's interest and needs. In recent times the globalisation of data and information, and especially environmental information, is widening but also the need to structure the overload of data according to user's needs and different practical purposes of data delivery. WISA is a combined information system which keeps in a data warehouse data, but links also to distributed data sources.

***Keywords**: Water Information System; E-Government; Web GIS*

1. INTRODUCTION

According to Article 59 of the Austrian Water Act 1959, the Water Information System Austria (WISA) has been implemented. WISA provides the public and the administration with for them relevant water data and information. It is designed as an on-line information system integrating already existing distributed data sources and distributed information systems as well as a data warehouse with aggregated data.

Following the legal regulations of the Austrian Water Act 1959 and the Environment Information Act, the access to environment data has become more important than it was ever before. Thereof it depends a lot how the data are accessible or presented and the aggregation of data should correspond with the user's interest. In recent times the globalisation of data and information, especially the transparency of environmental information, is widening but also the need to structure the overload of data according to user's needs and different practical purposes of data delivery. This document describes the system itself with focus on a combined approach of central data keeping and data access to decentralised databases. It is a combination of a central data warehouse, keeping aggregated data, and service-orientated architecture, linking the distributed water-related databases with access to disaggregated data.

2. DATA ACCESS IN WISA

Beside other requirements, the intention of the Austrian Water Act is to use WISA for the collection and usability of water data all over Austria. Therefore it shall serve as a tool to provide an overview of the whole water sector in Austria with relevant data. This means that aggregated data and pre-defined data are to be held in WISA. Further it is required to provide the user with background data, processed by several sectoral water databases which are linked to the WISA portal, see Figure 1 – WISA Portal.

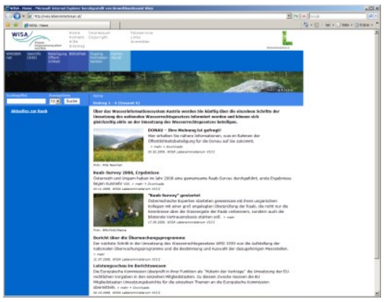


Figure 1

The information will be structured according to the delineation of international river basins and Austrian sub-river basins (national planning areas). On the other hand the information is gathered according to a thematic view and structure. WISA therefore helps to survey the status of the waters using the monitoring data of the national water quality monitoring programmes and gives characteristics of the river basins und sub-river basins. Besides that, it delivers information on significant pressures in the river basins.

The data and system shall be ready to be used also for the purpose of analysing the impact of human uses on the status of surface waters and groundwater, and to perform an economic analysis of water uses. In order to support the regional water managers in the process of developing a cost-efficient programme of measures for the national river basin management plan, it is intended to provide a catalogue of measures.

WISA serves also for the dissemination of water management information to the public. The data according to the Austrian Environmental Information Act are freely accessible, except for selected information which may harm the interests of the enterprises. These data and information are accessible only for special user groups of the administration also because of the provisions of the Austrian Data Protection Act which ensures the protection of personal data. Therefore there was the need to implement a differentiated access to the water data with different user groups as well. In addition to the Austrian Data Protection Act WISA have also to take into account the requirements of the Austrian E-Governance Initiative, which means to fulfil the E-Government standards and conventions like WAI conformity.

2.1 From a central data repository to a distributed data management

In the first implementation phase the main focus was to build a data warehouse with all the benefits of keeping selected data at a central location. Technically the architecture of WISA has been designed as a data warehouse in connection with the WISA web portal as a central access point to water-relevant information in Austria.

The aim of WISA is to integrate distributed data sources into a homogeneous data warehouse. This is accomplished by two technically different technologies:

- Integration of e-Government compliant user interfaces of sectoral data bases into WISA
- Implementation of Web feature services to allow retrieval of distributed data

The data warehouse stores relevant metadata, geographical information as well as aggregated water quality data. The necessary information is provided by sectoral thematic databases over well-defined interfaces (OGC compliant web feature services). The user interface of WISA has been designed to be compliant with the Austrian e-Government standards. It comprises a central query tool to access information of the WISA database as well as Web-GIS functionality.

Basically, the system was planned first and foremost as an instrument for experts. Following the requirements of the Austrian Water Act 1959 - namely to let the public actively participate in the implementation - corresponding functionalities and possibilities were created and are still foreseen and under development. Basically, the public is to be provided with an easy-to-use website which complies with the Austria-wide e-Government standards, as well as Web-GIS applications tailored to several individual user groups and electronic form sheets, e.g. to fulfil WAI conformity.

The WISA Web-GIS viewer, e.g. Figure 2, provides the user with pre-defined thematic maps within the framework of a WISA Web-GIS with the possibility to add some selected information. The Web-GIS viewer provides standard functionalities and additional search functions for geographical search, thematic search and for searches according to coordinates. Furthermore, the application provides the possibility to request attributes to the GIS layers. Special care has been taken of the access to the metadata of the implemented maps. The metadata profile of WISA has been delivered according to the requirements of ISO 19115 and will also be harmonised with the requirements of WISE and INSPIRE this year.

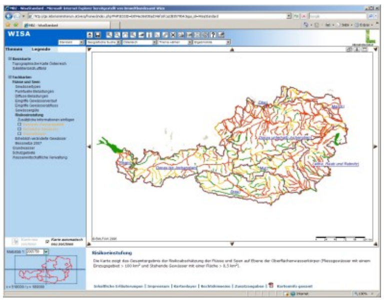


Figure 2

In 2008 a further expert version of the WISA Web-GIS viewer was launched, which allows the user to combine free selected layers and to store them as a self-created map via a user-specific "working place".

Furthermore an access to the Austrian water graph has been implemented, with several expert functions, which will enable the user to locate an object on the water graph and get the station (river kilometre and route-ID of the water graph, together with waterbody related information). In addition, the catchment area upstream or downstream can be queried from a freely chosen point on the surface water, which can answer many questions related to the analysis of impacts on water bodies and their catchments, see Figure 3.

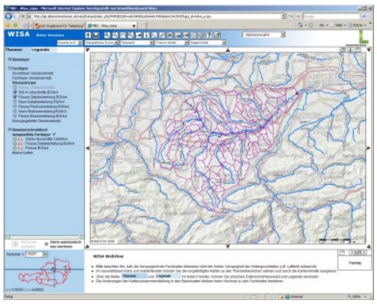


Figure 3

The expert user will be able to visualise e.g. monitoring sites, dischargers or hydromorphological structures in the selected catchment and query them. Such expert functions will be developed step by step especially for those users, dealing with authorisation and decision making processes in the water management.

So far data have been visualised mostly via GIS data which have been stored at the central WISA database. The next step will be the linkage of the objects in the GIS with the data of the sectoral databases and to have or to query aggregated data in data warehouse, too. So two different ways of bringing data to the user have to be mentioned: first, collecting data via interfaces from the sectoral databases and storing them in the central WISA database (mostly aggregated data) and second, routing the user to the sectoral databases, where the original data can be requested, see Figure 4.

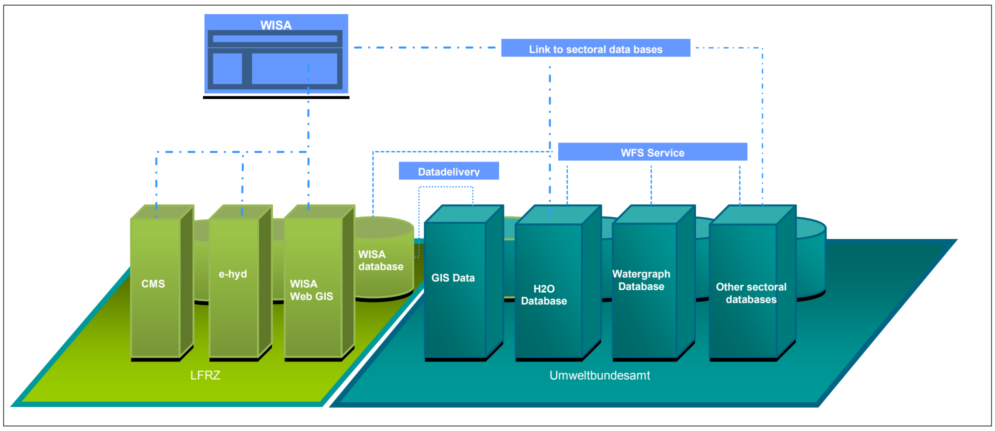


Figure 4

2.2 Web Feature Services - WFS

Web feature services at this stage of implementation are used to provide the WISA data warehouse with basic data from the sectoral databases. First the Austrian water graph with related information has been implemented in the interface and now its latest version can be requested at any time. Also, the interface for water quality data - on a yearly aggregation level of the data management system for Austrian water quality monitoring - has been implemented, and will enable the user to have information on each monitoring site for groundwater and surface water using the user interface of the WISA portal.

The Austrian water graph provides the necessary geographical information regarding surface waters: river segments, water bodies, catchment areas, etc., as well as domains (list of values) like bioregions, ecoregions, classifications, etc. The implemented web feature services comprise more than 50 features (with or without geometry). The feature “version” returns information on the version of the data set (major and minor release) and the underlying data structure. The interface provides the latest version of the Austrian water graph and supports the following calls: GetCapabilities, GetFeature and DescribeFeatureType.

The H2O Water Quality Database provides information about monitoring sites, monitored parameters and aggregated measuring values (number of values, maximum, minimum, arithmetic mean, percentiles) covering ground water and surface water. The implemented web feature services comprise version information, metadata and quality data, and approx. 20 features.

The implemented WFS operations comprise GetCapabilities, DescribeFeatureType, and GetFeature (no transactional operations have been implemented so far). Web feature services allow standardised requests based on Filter Encoding Implementation Specification) and deliver xml/gml documents as responses.

The software development environment of both interfaces comprises Apache TOMCAT 5.0.28 and Java 1.6. The underlying interface definition is Open GIS Web Feature Service version 1.0.

2.3 Integration of sectoral databases and the Austrian E-Governance

For a more detailed analysis, parts of the user interface of selected sectoral databases have been adapted and integrated into WISA, allowing access to disaggregated raw data. As the integrated modules rely on the same style sheets, the integration is seamless. The applications so far described are accessible for everyone and are considered to fulfil the requirements of free access to environmental information.

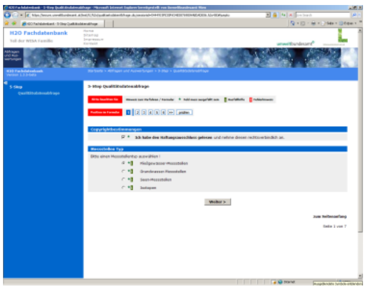


Figure 5

From 2007 to 2008 Umweltbundesamt successfully adapted part of the user interface of the existing H2O Water Quality Database and integrated it into WISA, see Figure 5. The user interface as well as the user / roles management have been changed to meet the requirements of the Austrian E-Governance Initiative (WAI conformity, distributed user rights management) and the style sheets for all water relevant web applications have been harmonised. As part of the infrastructure of e-Government Austria, authentication servers have been set up in all provinces providing the necessary authentication services to WISA and its integrated applications. The software development environment comprises Java 1.6, ECLIPSE 3.3 as IDE, Apache TOMCAT 5.0.28 and Apache Validator. Enterprise Architect is used for system specifications.

2.4 WISA Ontology

The aim of WISA Ontology is to allow overviews over datasets of the Austrian Water Information System, their origin, interfaces, targets and the connex with the reporting obligations of the Water Framework Directive. It is established in several steps, starting with basic datasets in 2006 / 2007 and now getting extended by remaining topics, offering an overview over the pathway of all WISA data from their investigation to the final report. We are using FreeMind (Ver 0.8.1 freeware) for the very intuitive design, Protégé (Ver 3.3, Stanford University, open source) (based on Protégé Server and the collaborative plug-in) for the formal description in OWL and Topbraid (ver .2.5.3, Topquadrant) for the import of files. The graphical display is done with Jambalaya (ver 2.6.1, University of Maryland, Protégé plug-in), which may be replaced, since it is not fully satisfactory, and does not allow sufficient drill – into.

The main concepts of the WISA Ontology are about the dataset, the object and parameter it is dealing with the source of the data and the structure behind it, the method how the parameter is won from the source, including details of the interfaces and referring to the reporting obligation.

2.5 Future Aspects

As a first step on the way to a more decentralised system it is intended to implement, in the expert version of the WISA Web-GIS, a "GIS-Explorer" which will function as a meta catalogue service and will support the user to search for geographic resources. The system shall use web services and find information sources in the web with water related topics. With this tool any service based geographic resource could be combined with information provided by WISA via the WISA Web-GIS.

Future extensions of WISA are planned in the following directions:

- Extension of the data coverage by integrating more sectoral databases
- Higher level services by implementing additional WFSs (including transactional WFS) and Web Mapping Services
- Extend the potential users of the implemented infrastructure
- Integration of WISA ontology

Extending the accessibility of water data step by step is one of the primary goals in the future. In some cases this will result in the organisation of data in new sectoral databases and their integration into WISA too.

Additional services are foreseen on the output side (improved accessibility of data) as well as on the input side (e.g. the use of transactional services for data assembly). As far as data assembly is concerned it would be desirable to integrate the Austrian provinces as main data sources. Data deliveries into the thematic nationwide databases could be accomplished by services, see Figure 6.

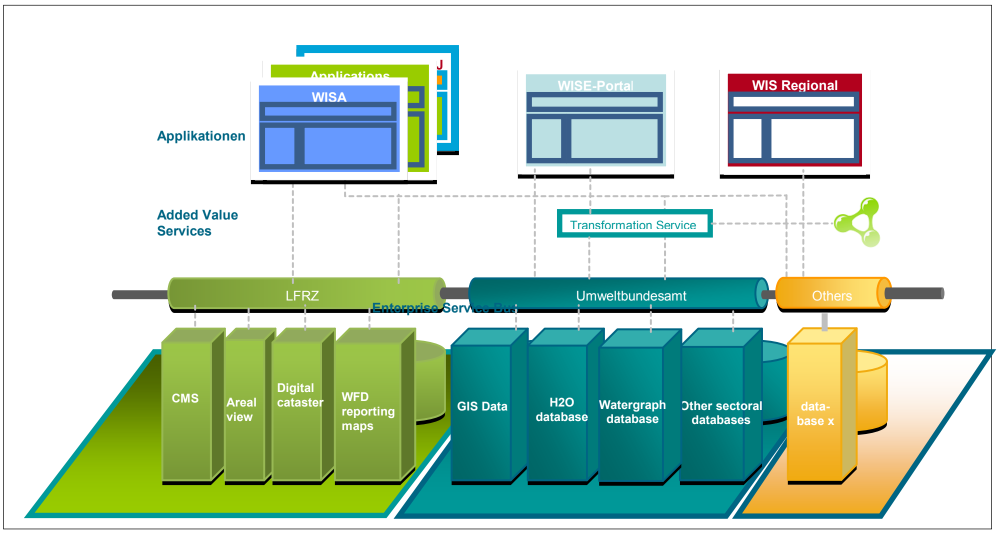


Figure 6

3. CONCLUSIONS

It is important to have selected data stored at a central point in a rather static form, especially for the purpose of keeping and versioning the data subject of reporting. If users wish to get information on the current state of e.g. water quality nearby their home, it is rather more appropriate to query the operative sectoral data bases, than a data pool with data updated every three years. Both ways of data keeping and data access make sense and are important. WISA tries to combine both approaches and to provide requested information to the user in a user-friendly and most appropriate way.

REFERENCES

WISA – Water Information System Austria: http://wisa.lebensministerium.at

Wasserrechtsgesetznovelle (BGBl. I Nr. 123/2006). Amendment to Austrian Water Act 1959 (Wasserrechtsgesetznovelle 2006)

The new monitoring portal: sharing environmental monitoring information in the Netherlands.

J.G. Nienhuis

Netherlands Environmental Assessment Agency (PBL), pb 303, 3720 AH Bilthoven, The Netherlands (info@pbl.nl)

Abstract: In the Netherlands, many agencies are involved in monitoring of the environment. National institutes, provinces and NGO's are involved. Monitoring data is used by the Dutch provinces, the Dutch Government, as well as by environmental agencies in Europe. There is a lot of information available (Albers, 2007). But the problem is how to find specific data. A new environmental monitoring portal has been built, to make finding this sort of information easy and efficient. It does not contain any information itself, but links sources (contact information on people and organisations) to environmental topics, websites and to the data on websites. The site is meant for all who need information on environmental monitoring in the Netherlands. They can be researchers, policymakers and advisors of the Dutch Government and the provinces, but also people from national water management boards, city councils and research institutes. The portal covers all the environmental issues: from waste management to biodiversity. The website is updated weekly. The aim is to present information on the monitoring of the environment: 'Who is doing what?', 'Who do I contact to obtain specific environmental data?' and 'Where can I obtain information on a specific environmental topic?'. Visitors can search for information by theme, by source of knowledge, or by entering a key word. This enables very direct searches, which is practical and saves visitors a lot of research time. The new website is made by the Association of the Provinces of the Netherlands (IPO), as a joined initiative and with participation of the Netherlands Environmental Assessment Agency (PBL), the National Institute for Public Health and the Environment (RIVM), the Dutch Ministry of Housing, Spatial Planning and the Environment (VROM), the Dutch Ministry of Agriculture, Nature and Food Quality (LNV) and the Dutch Ministry of Transport, Public Works and Water Management (V&W). Harmonization of information of provinces and national monitoring institutes, still poses many problems. Most of these problems concern indicator definition, level of experience and difference in reliability of data. This portal offers its visitors the correct information and a possible solution to these problems, by enabling them to find and contact the right people and organizations that collect or use monitoring information.

Keywords: Monitoring; Portal; Internet; Environmental data; Environmental information.

1. INTRODUCTION

Although the Netherlands is a relatively small country, there are many agencies involved in the monitoring of the environment, nature and water. Besides the Dutch Government and local and regional authorities, the various environmental topics are also dealt with by national research centres and institutes, Dutch water boards, NGOs and others. Monitoring data is used in many separate monitoring programmes, and is also shared with other institutes, such as environmental agencies in Europe. And although in most cases formal agreements for sharing data are very well organised, finding the right people, projects, activities, data collections and organisations still remains difficult.

There are many questions to answer: who is doing what; which person of which organisation is working in a certain field of interest; which regulations are the most important for my topic; which reports and papers are written on a certain topic; where can I find the data I need quickly and who can tell me more about it?

To answer all these questions, a new environmental monitoring portal has been built. In this paper, the portal is introduced. It shows which information the portal currently contains and what the plans are for the future. A few early lessons learned are elaborated and, subsequently, a number of interesting possibilities for other EU countries will be looked at.

2. WHAT IS THE MONITORING PORTAL?

The monitoring portal is a website made by the Association of the Provinces of the Netherlands (IPO). The website is a joined initiative of several governmental organisations: IPO, the Netherlands Environmental Assessment Agency (PBL), the National Institute for Public Health and the Environment (RIVM), the Dutch Ministry of Housing, Spatial Planning and the Environment (VROM), the Dutch Ministry of Agriculture, Nature and Food Quality (LNV) and the Dutch Ministry of Transport, Public Works and Water Management (V&W). It contains data concerning the environment in the Netherlands and offers easy access for people and organisations. It is used, for instance, for fulfilling monitoring obligations, for carrying out monitoring activities, for papers and websites.
The portal provides the links to the data; mostly, it does not contain the data itself. Only in cases where the internet does not provide access to the data, it is stored on the portal website itself. The portal's goal, however, is to provide easy access to data through links, rather than storing any data itself.

An easy-to-use search engine has been built into the portal, to give people fast and easy access to the data.

Moreover, to make the portal and its links easy to find, special care is given to the sharing of information with the Google search engine.

The monitoring portal was built by combining the best parts of two websites that had already existed for several years. One of the websites provided information on the work of IPO and RIVM on several environmental topics (www.iporivm.nl). The other website held information on the obligations and activities concerning the monitoring of the environment (www.milieurapportage.nl). Both websites had proven successful over the years, but by combining their information and by making sure it can be accessed easily, the portal is believed to provide the information in a way that is even better, faster and easier to use.

The portal stores information about:

1) People. Who are working in the monitoring field, what is their expertise and how can they be contacted?
2) Organisations. All organisations working on monitoring topics. What they do, but also how they can be reached, including possible links to their websites.

3) Obligations and activities in the field of monitoring. What is needed to fulfil all the obligations of the Dutch government agencies? What is done by which agency or institute and how can I contact them?
4) The monitoring results. Links to data, papers, websites and systems, resulting from all the activities above.

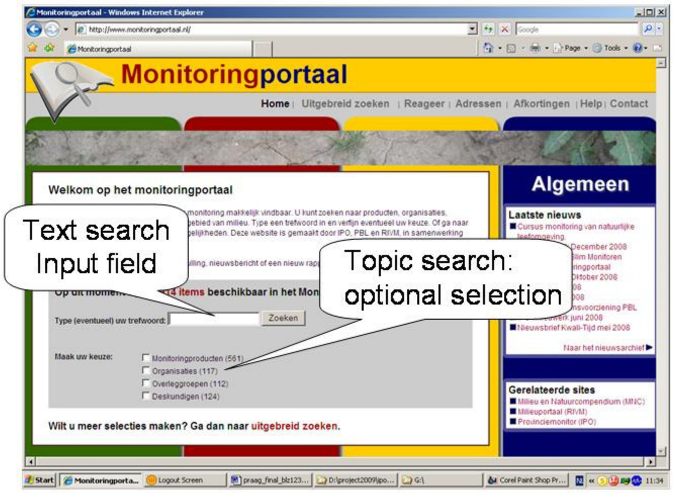


Figure 1: The website [www.monitoringportaal.nl.](http://www.monitoringportaal.nl) People can start looking for data using text searches or by selecting a specific topic.

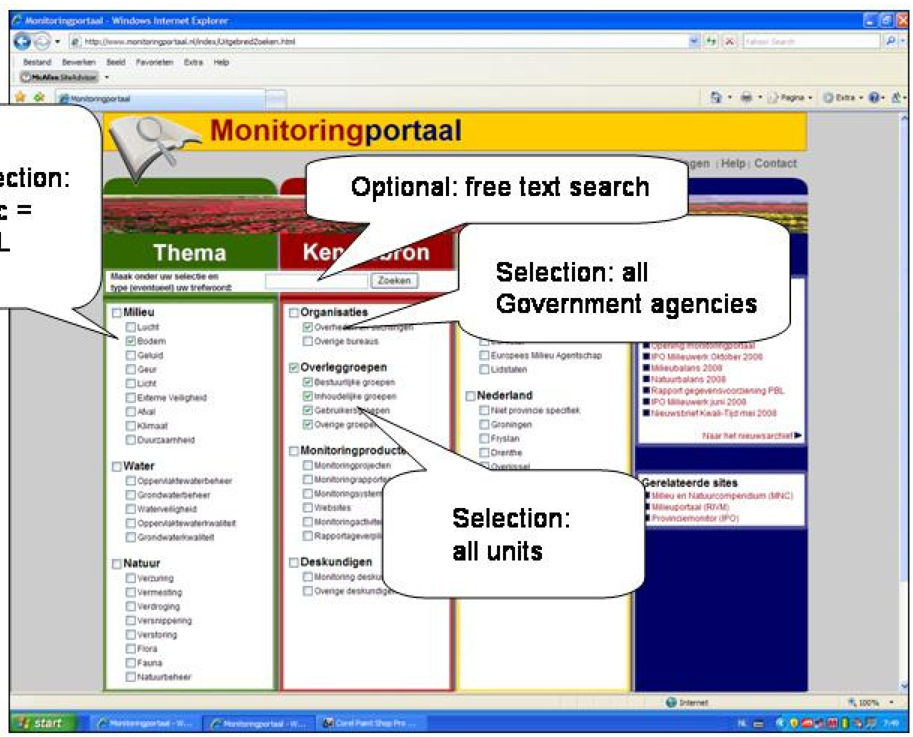


Figure 2: A few examples of the advanced search options providing access to all the data and cross links in the monitoring portal. This example results in the links to 81 items: 30 government agencies and 51 units and groups of people, working in the field of monitoring the topic “soil”.

All the data is cross-linked within the portal, so that when you find a person working in a specific field, you also find information on his or her organisation and how to contact them. The portal operates with an easy-to-use and flexible graphical user interface, which enables visitors to find what they are looking for quickly and efficiently.

Special care is given to the maintenance of the information on the portal itself. This data has to be correct and current. A team of editors check, update and promote the portal, so more and improved information is put on the portal continually. Portal visitors can also provide information directly to the editorial team by filling in a simple online form. The editorial team uses a specially made Content Management System (CMS) for flexible and efficient maintenance of the portal. The website is updated weekly. Old items are removed, corrections are made and new information is put into the system.

Initially, the portal only provided information on the monitoring of the environment, but now it also contains information on other topics, such as the monitoring of nature and water. In the future, even more topics could be added.

3. LESSONS LEARNED

The new monitoring portal was completed in October 2008. In 2009, a full evaluation will be conducted, but some interesting first lessons can be learned already.

Comparing the number of visitors of the two old sites with those of the new site, is showing a very promising result: the new portal has attracted almost twice as many visitors in its first few months.

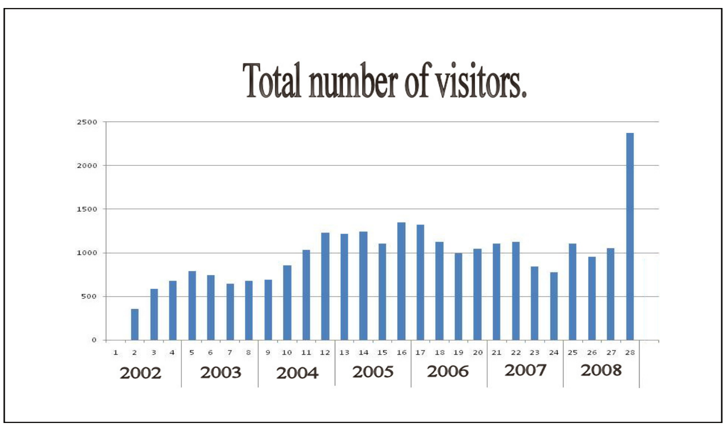


Figure 3. Number of visitors: the effect of the new portal, starting October 2008. The data is giving for 7 years, starting with the 2 original sites in 2002.

Monitoring the use of the monitoring portal is conducted in several ways. The number of visitors is stored, as shown in the previous figure. In addition, the way people use a search engine (Google) to visit the portal is monitored carefully. Each series of search words that is entered in Google is shown and stored. Looking at these search words, some conclusions can be drawn. It is interesting to see that the portal is visited by all sorts of organisations within the Netherlands; all the major national institutes have visited the portal in recent months, as well as most of the other institutes and organisations working in the field of environmental monitoring. Most of these searches were aimed at finding other people. Who is doing what? How do I reach a certain person? I know the name, but I have no contact information… Other searches are conducted to get information on a specific environmental subject; the details about a specific pesticide in the soil, for example. The monitoring portal gives the complete list of what is monitored in soil, and gives the names of the organisations working in that specific field.

So, looking at the first results, we conclude that the portal has made a successful start. A complete evaluation will be conducted at the end of 2009, but the first results are very promising.

4. OPPORTUNITIES FOR OTHER EU COUNTRIES.

The monitoring portal is a useful source of information on monitoring within the Netherlands. This approach of providing access to this sort of information could work, in a similar way, for other EU countries, too. The way the portal is built, the information it contains and the lessons learned could all be a source of inspiration for other countries.

Perhaps you are also thinking of setting up a similar system for sharing information with all people working in the field of monitoring. A lot of activities are conducted in cross-border projects and agreements. The reasons for this are easy to understand. Many environmental problems do not stop at national borders and many people of different countries are already working together on these topics.

The content and the user interface of this first version of the portal is in Dutch only. To provide the information of the monitoring portal for other EU countries it might be a good idea to convert the text of the portal, using automated text conversion tools like Google Translate.

In several other EU countries websites to provide access to monitoring information are available too. The best thing to do now, is to start finding ways to combine these efforts. Perhaps in the SEIS context, we can combine websites, portals, data and lessons learned, to provide this information for all the people looking for monitoring information in the EU.

REFERENCES

Albers, R. Towards effective Monitoring; main report for Environment, Nature and Water. *Seminar 'Towards Effective Monitoring', Oct 25&26 2007 in Brussels*

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Directive INSPIRE transposition in Bulgaria

Kristian Milenov and Ivan Philipov
Agency for Sustainable Development and Eurointegration (ASDE), Bulgaria
(asde@online.bg)

Abstract: Since 2004, efforts have been made to structure the Bulgarian SDI activities. Despite Bulgaria is a new MS and the Directive 2007/02.EU - INSPIRE is the first one, officially published in Bulgarian language, the Bulgarian administration, supported actively by the civil society sector fulfilled the required actions. The actions developed till now are coordinated by two agencies and they are playing an important role in process of implementation as partners in a public private partnership (PPP) way: The State Agency for Information Technology and Communications (SAITC) which is the state structure, responsible by resolution of the Council of Ministers for the transposition of the directive, and The Agency for Sustainable Development and Eurointegration (ASDE) – which is an NGO, for public purposes, acting as initiator and partner in the process. The objective of the partnership is to coordinate activities of different ministries, specific state agencies, regional and local authorities producers and users of spatial data. The collaboration with the private sector, scientific and academic organizations is also seen as an important precondition for success. This approach was examined by joint work on a project for risk management through Geomatics, especially for earthquake risk assessment and flood risk prevention, and their integrated risk effect. Some examples are illustrated. We believe that our good practices could be successfully implemented in realizing the priorities of GMES/Kopernikus including the close relation between GMES and INSPIRE.

Keywords: PPP; GMES; INSPIRE; risk management.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

An Introduction to the Aggregation Database

Søren Roug

European Environment Agency, Kongens Nytorv 6, DK-1050 Copenhagen K, Denmark
(soren.roug@eea.europa.eu)

Abstract: Reportnet is the European Environment Agency's tool to collect environmental reporting from the member states. In practice it is a pipeline consisting of delivery, quality assessment, aggregation and analysis. Until recently the main streamlining effort has been on the first two segments. The providers know what to deliver, when and in what format. The requesters get automatic quality analysis and a feedback channel to the provider. This paper will describe a possible solution towards the aggregation step in the pipeline. It is called the aggregation database and follows SEIS principles inasmuch as it discovers and harvests datasets on remote systems and aggregates the content in a database while keeping track of the sources. The same mechanism makes gap-filling of the merged deliveries less like manipulation, because the filling of the gaps is created as just another source that is part of the aggregation and can be traced.

Keywords: Aggregation, gap-filling, SEIS user-interface, discovery service, traceability.

1 PURPOSE IN REPORTNET

The aggregation database is planned to be an extension of Reportnet's search engine, the socalled Content Registry (CR). It is going to serve two purposes:

1. To discover and keep track of SEIS datasets. It does not matter what format these datasets are. MS-Excel, HTML etc. are all equal.
2. To understand those SEIS datasets that are in XML format. It will do this by transforming the content to a quadruples structure and storing it in a database.

The system is not going to be the only way to get access to the Reportnet datasets, but to be an extra tool in the toolbox.

2 DISCOVERING THE DATASET

Datasets tend to be only numbers. Unlike Google, that mainly harvests webpages, a SEIS search engine can't get information from the content of the files. The aggregation database will use two mechanisms to discover relevant datasets.

1. A collaborating node provides a *manifest* of the relevant files existing on its own website. The manifest lists location and any metadata the provider has on each file. This manifest must be downloadable over the Internet and will be periodically harvested by Reportnet's CR. When the file occurs on the manifest, it is then discovered by Reportnet.
2. The user finds the file with Google, surfing the web, or is owner of the dataset. He *registers* it at Reportnet's QA Workbench (QAW). This site is for the registration purpose nothing more than a special collaborating SEIS node that provides a manifest to the CR.

Once the content registry has discovered the existence of a dataset, the users can *amend* the available information on the file by adding metadata to it in QAW. There are several relevant pieces of metadata to enter; title, tags, coverage – both spatial and temporal, what purpose the dataset was created for, what methods were used, its quality etc. All these can be entered at the QAW. But a glowing review of a dataset can only be trusted if we know who wrote it. As we'll see, QAW has metadata on the metadata.

2.1 What sites will have a manifest file?

As many as possible. Relevant datasets aren't just the ones that the member countries provide. We want to be able to combine data. If we have a list of airport locations and how many flights they have a year – that's relevant. Dams and their locations. Basic statistical data on countries – number of inhabitants, size, GDP – is needed to understand other data. Eionets database of reporting obligations is a database, but every obligation has a factsheet page, and the obligation *looks* like a file. That database should have a manifest because Reportnet deliveries declare in their metadata they are responses to obligations (implying a relationship).

Manifest files don't have to follow just one format. There is room for integration with the INSPIRE discovery service, and there is room for the INSPIRE data formats.

3 TRACKING THE DATASET

Once a dataset has been discovered by Reportnet's CR, it starts to check periodically whether it is still at the remote location, and whether it has been modified. We are planning to check every 6 weeks. If a dataset is found to be unavailable its record in the search engine is marked as stale, and a notification is sent out via the Unified Notification Service (UNS) to whoever is subscribed.

3.1 Examples of tracking

ROD contains references to repositories, legal instrument documents and guidelines. Until now we have used special purpose-built software to check the links and email an administrator when a link went dead. With the new way to track, a simple feature will be added to ROD that provides a manifest of remote references. The CR will automatically check the links, and all the administrator has to do is to subscribe to UNS for dead links in ROD.

Precisely the same mechanism can be used in many other contexts. One is the SERIS database at http://www.eionet.europa.eu/seris.

4 USING THE DATASET

All the registration and tracking isn't of much use unless the users can find the datasets again in CR. The plan is of course to use the metadata to search for the datasets. The user types in 'soil' and gets what's known. The usefulness of the system hangs on getting as much metadata as possible. We will therefore encourage people to 'improve' the metainformation as much as possible. E.g. When the user has successfully found some relevant datasets we'll make it easy to enter more metadata by tagging all the hits with one tag.

The user has found some datasets – then what? In Google he can click to the webpage. He'll be able to do the same here. But we're also considering keeping a copy of the dataset when the CR first discovers it, because it will make it possible for us to provide the dataset in other formats. The user might not be able to use a Shapefile. We can make it viewable as an image.

5 UNDERSTANDING THE DATASET

If the dataset is in a known XML format, CR will be able to understand the *content* of the dataset. A file with a known XML format is one that has a schema identifier; or has a persistent location on the Internet, where the XML format has been discovered in some other way. The way to understand the content is to *transform* it and put it into a database.

Let's imagine that Austria reports two new Airbase stations. Number 32301 and 32302 They could store the information in an XML file called *stations.xml* like this:

```
<stations xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="http://air-
climate.eionet.europa.eu/o3excess/stations.xsd">
  <station>
    <local_code>32301</local_code>
    <name>St. Pölten - Eybnerstraße</name>
    <longitude_decimal>15.63166</longitude_decimal>
    <latitude_decimal>48.21139</latitude_decimal>
    <altitude>270</altitude>
    <station_type_eoi>Industrial</station_type_eoi>
    <area_type_eoi>urban</area_type_eoi>
  </station>
 ...
</stations>
```

It will be simpler to explain if we show the data as a table like shown below. The transformation will take any XML content and store it in a database table with only four columns called Subject, Predicate, Object and Source. Every record gets a *type*. This one is called "Station".

Station structure as a table

Identifier	local_code	name	...
#32301	32301	St. Pölten - Eybnerstraße	...
#32302	32302	Europaplatz	...

Quadruples

Subject	Predicate	Object	*Source*
#32301	type	Station	*stations.xml*
#32301	local_code	32301	*stations.xml*
#32301	name	St. Pölten - Eybnerstraße	*stations.xml*
#32302	type	Station	*stations.xml*
#32302	local_code	32302	*stations.xml*
#32302	name	Europaplatz	*stations.xml*

Now imagine what will happen when CR loads files of the same format from several locations. The data is automatically aggregated. Hence the term aggregation database. You can even load files with conflicting information on the same stations. They are kept separate because the Source values will be different.

The mechanism makes it possible to have other sources add columns (predicates) to a table if the key (subject) is the same.

5.1 Querying the database

The question is how do we make use of it? Here we must implement a query language in CR. If we start by something simple, such as a query on all records of type Station, CR will then find all rows with the predicate=”type” and object=”Station”. It would then for all the subjects it has found look up which predicates exist and show a table.

Identifier	Local code	Name	...
#32301	32301	St. Pölten - Eybnerstraße	...
#32302	32302	Europaplatz	...
...	...	...	...

The data could then be exported as MS-Excel, MS-Access or whatever else is convenient for the user.

6 TRUSTING THE DATASET

How would you know that a dataset you’ve found is trustworthy? As always, you look at the website that hosts it. Is the site to be trusted? But for datasets, there is more. Is it produced for a purpose compatible with your own needs? Is it of sufficient quality? The QAW website will store all the known metadata, and when you know the URL of a dataset you can query the QAW for the metadata. But QAW will also store where the metadata came from. For example if someone has registered *stations.xml* with QAW and entered information on the methodology, QAW will store which user wrote it.

But how do you trust aggregated data? The same way. By inspecting the source. But since it is aggregated, there must be a source on *every* field. You will be able to look at the data separated, or *overlapped*, in a way that new data overwrites old data for the same identifier. A list of overlapped records could look like this, using colour codes to show the source of the fields:

Identifier	EoI	Updated	Local code	Name	Long	Lat
#32301	AT44	2008-05-28	32301	St. Pölten - Mühlweg	15.62880	48.21194
#32302	AT220	2007-05-31	32302	St. Pölten - Europaplatz	15.61111	48.20999
...		...	...	...	...	

	Source: http://www.uba.at/stations.xml	[Turn off]
	Source: http://air-climate.eionet.europa.eu/airbase-stations.xml	[Turn off]
	Source: http://www.uba.at/station-updates.xml	[Turn off]

You will be able to inspect the metainformation for each source, turn individual sources on and off and when satisfied, export the data as MS-Excel, MS-Access etc.

7 GAPFILLING

You can weed out bad data by excluding the source, but what if you want to *correct* some data? This can be achieved by creating your own source. There are two ways; copying an existing source or creating an empty. In both cases, the source will be stored on QAW.

Identifier	Local code	Name	Date	Longitude
#32301	32301	St. Pölten - Eybnerstraße	2005-10-8	15.63166
#32301		St. Pölten - Mühlweg	2008-6-18	15.62880
#30299	30299	Gent	2004-11-12	3.733333
#42882	42882	Köln	2001-4-14	6.966667
#32301			2008-11-27	15.63100

Your gap-filling layer

What you normally will see is that you will be able to change any value in the table, and the change will be stored in your gapfilling layer. Others will see your layer as yet another source to consider for the merge.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pilmmann, I.Holoubek, R. Legat (eds.)
Masaryk University, 2009

NESIS - a Network to enhance a European Environmental Shared and Interoperable Information System

G. Saio[1], **D. Vandenbroucke**[2], **R. Wawer**[3]

[1] Geographical Information Systems International Group-GISIG (g.saio@gisig.it)

[2] K.U.LeuvenR&D Division SADL-Spatial Application Division (Danny.vandenbroucke@sadl.kuleuven.be)

[3] K.U.LeuvenR&D Division SADL-Spatial Application Division (Rafal.Wawer@sadl.kuleuven.be)

Abstract: In a framework oriented by the SEIS communication, the NESIS Network aims to promote the uptake of ICT solutions to address the basic problems faced by public authorities in providing information related to monitoring and reporting environmental impacts and threats.
The Network targets to leverage the stakeholders' community, starting from EIONET and its experience to develop a widely agreed action plan for SEIS as a distributed, standards-based infrastructure for spatial and non-spatial environmental information, based on the principles of shared access rather than centralized reporting.
NESIS analyses the current State of Play of environmental information systems and services in Europe and collects Best Practices, to finally derive and discuss an ICT roadmap for SEIS implementation

The work plan is organised along two years and a half: collection and analysis of practices on environmental information management, definition of an ICT roadmap for SEIS implementation, validation of it by the partners and a group of stakeholders from different Countries, preparation of the network follow-up, with additional involvement of the stakeholders and an impact assessment of the roadmap at national level.

The expected impact is the creation of a collaborative framework to share Best Practice on ICT for environmental data management and on the related procedures for an effective use of the SEIS Framework. In the process of SEIS implementation NESIS intends to accompany the process as "human and organisation" component.

The project has performed in its first phase a survey about the State of Play in management environmental data at national level, following the methodology of the INSPIRE State of Play approach, as well as has started collecting "good practices" of ICT solutions for managing environmental data.

Keywords: Network; SEIS; ICT roadmap; Best Practices; stakeholders' involvement

1. INTRODUCTION: NESIS AND ITS CONTEXT

A communication for a Shared Environmental Information Systems (SEIS) has been released on 1st February 2008, to draft the rules for:

- improving availability and quality of information needed to design and implement Community environment policy;
- reducing administrative burden on Member States and EU institutions & modernise reporting;
- fostering the development of information services and applications that all of us can use and profit from,

and with the scope of an integrated and sustained environmental information system able to improve the sharing of data and to lead to an upgrade, both in the quality of environmental data and information, and in its management, use and dissemination.

In this framework, the NESIS Network targets to promote the uptake of ICT solutions to address the basic problems faced by public authorities in providing information related to monitoring and reporting environmental impacts and threats.

The management of environmental information evolves towards a distributed, standards-based infrastructure for spatial and non-spatial environmental information, based on the principles of shared access rather than centralized reporting. This matches with the latest developments in the ICT domain and in terms of the evolving legal framework, including the implications of the INSPIRE Directive.

In supporting sustainable growth, specifically enhancing environmental, the NESIS Network is expected to bridge the gap between the ICT domain and the public authorities mandated to create, manage and exchange environmental information.

The NESIS Network targets to leverage the existing community of stakeholders starting from EIONET, a network of some 900 experts from over 300 national environment agencies and other bodies dealing with environmental information in 38 European Countries. By reviewing the state of play of current ICT practices for monitoring and reporting within the EIONET community and the requirements to enhance the existing infrastructure ICT deployments, the NESIS Network intends to support a widely agreed action plan for SEIS.

The outcomes of the NESIS Network, even if targeted at EU wide policy requirements, are focused to foster the interchange of information between regional, national and international authorities, as opposed to the flow of information from the Countries to the EU, which typifies today's reporting methodologies.

2. NESIS DEVELOPMENT

NESIS analyses and describes the current State of Play of environmental information systems and services in Europe and prepares and discusses a ICT roadmap for SEIS implementation and collect information on good practices on ICT solutions for managing environmental information.

The work plan is organised along two and a half years. The first one is mainly devoted to the national level with the collection of practices on environmental information management by the EIONET National Focal Points (NFP's), to provide coherent "good practice" inputs. An analysis will be prepared based on the collected material.

The second year is dedicated to define a ICT roadmap for the SEIS implementation and such roadmap represents one of the most relevant project outcomes. It is prepared taking into consideration the report of the status of environmental monitoring and reporting in Europe, and is validated by the partners and a group of stakeholders from different Countries. The roadmap will be discussed with DG INFSO, the EEA and the JRC, for validation and to start the process of public consultation in view of the adoption by the Member States.

The final steps of the project are instead devoted to the preparation of the network follow-up, with additional involvement of the stakeholders and an impact assessment of the roadmap at national level. In this period the NESIS exploitation policies are also defined, so to be aligned with the SEIS implementation process and its operational adoption.

The networking approach implies the sharing of best practice as a process running during the whole project and as an instrument to assure the involvement of the targeted group of stakeholders. They are expected to provide and share with the other members their own best practice, in view of the future sustainability of the NESIS open and enlarged network.

The NESIS expected impact is the creation of a collaborative framework to share Best Practice on ICT for environmental data management and on the related procedures for an effective use of the SEIS Framework throughout the involvement of a wide number of stakeholders. Starting from the definition of the implementation roadmap the long-term perspective is a wide sharing of experience on the test and the use of SEIS. In such way NESIS intends to accompany the process of SEIS implementation as its “human and organisation” component.

The NESIS Network will be also proposed as a SDIC (Spatial Data Interest Community) to support the implementation of INSPIRE.

3. THE FIRST STEPS OF NESIS

The first phase of the project has been dedicated to two main tasks:

- the activity at national level for the analysis of the State of Play about the ICT components that will contribute to the development of SEIS in the different Partner Countries, according to a common analysis and reporting template
- the identification of good practices of ICT solutions for managing environmental data and collection of relevant information about the above practices according to a template, as well as the organisation of the collected practices into the “NESIS Best Practice Catalogue”

3.1. THE SURVEY ON THE STATE OF PLAY

The project, which enjoys also the opportunity of the SEIS Country Visits (for a better contact with the stakeholders), as well as of sharing their outcome, performs as aforesaid in its first phase a survey about the State of Play in management environmental data at national level. The survey’s background and methodology lies in the INSPIRE State of Play approach (Vandenbroucke, Janssen, 2008).

The survey is structured upon:

- organisation of the technical work
- overview of SEIS related activities
- data and information flows
- infrastructure and systems
- use of standards
- obstacles
- priorities for SEIS
- practices

The survey is implemented throughout an extensive template, meant to provide a concise summary regarding the status of the ICT components that are already in place in member states and could contribute to SEIS. It indicates also those components that should be considered as good or best practice in the field.

The section addressing organisation of the technical work does not aim at repeating what has already been described in the EEA SEIS country report. It rather complements this

information focusing on organizational set-ups for following-up and coordinating technical issues. Findings that should arrive from this section are for example related to:
- organisational set-up for technical issues
- technical implementation plan(s) and guidelines
- funding

The overview of SEIS related activities highlights specific technical activities related to SEIS, as well as activities that aim to learn/test certain technical solutions. Information collected within this section are for example related to:
- User requirements analysis
- Review and evaluation of monitoring & reporting mechanisms
- Pilot projects
- Link with INSPIRE/GMES activities

The section data and information flows focuses on the data/information of SEIS, including the metadata. It relates to Geographical information, statistical data, in situ data, Earth Observation data, etc. Findings that should arrive from this section are for example related to:
- Electronic versus analogue data
- Major information flows
- Data sources and metadata
- Streamlining monitoring and reporting
- Data sharing

Infrastructure and systems section focuses the technical infrastructure, and applications developed. Findings that should arrive from this section are for example related to:

- SEIS architecture
- Existing systems and applications
- Network services
- Portals
- Use of Reportnet

The utilization of standards in information exchange is recognized as a fundamental requirement towards interoperability of data and services (Crompenvoets J. et al., 2008). This section should give an overview of which standards are used and how they are implemented (e.g. W3C, ISO, CEN, ETSI, OGC, IETF, OASIS, …). This includes 'de facto' as well as 'de jure' standards, international as well as national, GI as well as general ICT and community specific standards. Standards are important to reach semantic as well as technical interoperability. Findings that should arrive from this section are for example related to:

- Generic ICT standards
- GI and EO related standards
- Community specific standards

An important part of the survey is related to possible obstacles for the implementation of SEIS in the member states. The NFPs describe the major technical difficulties encountered as well as problems related to the organizational set-up for follow-up of technical issues. They describe also briefly how such the revealed obstacles could be overcome . Having the overall view on the obstacles and current state of the use of standards, infrastructure and data flows, NFPs are asked to allocate funding for SEIS implementation components in an order of priority in their country, giving the total funding amount of 500.000 Euro. This section should reveal priorities of SEIS implementation in Member States participating in the survey.

The last section of the survey addresses good practises in sharing environmental information. The section is meant as a explanatory add-on to the NESIS Catalogue of Best Practise, giving the opportunity to explain and motivate why NFPs think a component that is (being) developed in the framework of SEIS is an example of a good practice, and that could be of help for preparing the SEIS roadmap for Europe. Good practices can be particular for NFP's set-up, but are preferably useable in other European countries. NFPs were encouraged to indicate also good organizational practices.

Preliminary results of the State-of-Play

The survey template was subject to consultation and review with the Steering Committee and the Advisory Board, and with JRC and EEA in order to be complementary with SEIS country visit report for dealing with ICT aspects of SEIS.

Ten national SoPs have been drafted, provided by NFPs representing from the following Countries:

Czech Republic	Iceland	France
Norway	Austria	Sweden
Hungary	Malta	Lithuania
Italy		

The results from the SEIS National State of Play survey reveal considerable variability (among the questioned countries) in the needs and then in the definition of priorities for developing components of SEIS.

Although the SEIS implementation principles, expressed in the EC Communication on SEIS (EC, 2008), does not put much pressure on the data production, in five countries: Czech Republic, Hungary, Iceland, Italy and Malta; the activity is considered to play an important role in SEIS implementation. In case of France, where two scenarios of SEIS implementation were considered:

1. data is produced, collected and maintained by private sector
2. whole data life cycle is controlled and performed by national administration;

the first scenario predicts the acquisition of new data may play most significant role in implementation costs.

The most common need reflects the present lack of harmonization of existing datasets and information and the need for reduction of effort needed to fulfill monitoring and reporting obligations within EIONET

An important issue remains the lack of proper description of data. The implementation of INSPIRE Directive is already increasing the creation of metadata, but its influence is limited to geographical data, which is only a fraction of existing environmental data flows.

Based upon the hitherto survey results we depicted the most important SEIS implementation priorities to be considered on the European level (see Table 1).

	Activity	**Share [%]**
1	Creation of new data sets (GI, statistics, in situ measurement data, …)	9.1
2	Creation of metadata for data sets	9.8
3	Harmonisation of existing data sets and information	15.6
4	Re-engineering data/information flows in order to reduce the timeframe for monitoring and reporting	14.3
5	Automated creation of indicators and reports	7.6
6	Set-up of own network services to have easier access to the information	9.9
7	Set-up of network services at European level to have easier access to the information in a cross border context	7.1
8	Development of end-user applications	7.2
9	Implementation of standards	4.4
10	Set-up of better coordinating mechanisms	7.8
11	Incentives for better sharing	6
12	Other: coordination group	1.2
Total		**100**

Table 1. SEIS National State of Play. SEIS priorities.

The activities highlighted with light grey should be considered with the higher priority, with harmonization of existing datasets and information as well as re-engineering of data/information flows at the first places. They are related to organization and sharing of data and information on environmental monitoring at the level of national administration.

The activities highlighted in dark grey could be considered as less urgent, however they are important for dissemination of the environmental information and reporting as well as for assuring the access to information for the society.

The survey will be subject to annual updates and further national SEIS State of Play reports as well as summary SEIS State of Play at European level will be provided by NESIS network.

An important step in the assessment of existing SEIS components will be the publication of SEIS Directive, expected to be published in March 2009. NESIS will incorporate the legal framework into its annual survey on SEIS State of Play and fine tune the analysis on SEIS implementation priorities.

3.2 IDENTIFICATION AND COLLECTION OF BEST PRACTICE

This second task started in the first phase of the project deals with the identification of ICT solutions to manage environmental data and with a report of good practices as regards system infrastructure, dataflow and related services.

It allows the implementation of a “NESIS Best Practice Catalogue”, a living database updated/upgraded during the project development.

The Catalogue lets available meaningful technical information about good practices of ICT solutions for managing environmental data and is intended as a tool to exchange information across Countries and then to promote the participation of a wide audience of stakeholders at European level wishing to share their own experience and to query it to look for the experience of other stakeholders.

The Catalogue has been prepared via a factsheet structured upon a template, for uniform and reasoned collection of the needed information.

The template is structured upon the following sections:

- practice description,
- key-words to “characterise” the Best Practice,
- addressed environmental theme(s),
- geographical span (from EU to local),
- objectives and actors/users involved,
- processed information layers,
- data, metadata and information flow,
- architectural and technical aspects with the services set-up,
- status of implementation and planned development,
- lessons learnt,
- the references for further info.

After a first round the partners have identified a first set of good practices related to different environmental fields, Such “good practices ” (15 have been collected so far and have been described according to the template) are already available on the NESIS on-line Catalogue, accessible upon registration to the NESIS Network.

The following table shows the good practice collected in the different Countries participating in the network.

AUSTRIA	Environment Portal Austria
CZECHIA	Information System of Statistics and Reporting ISSaR Information System of Technical Environment Protection (ISTOZP) Czech Environmental Metaportal

FRANCE	Geoportal Environmental Portal National Interest Group on the Soils NatureFrance , portal of the SINP : Système d'Information sur la Nature et les Paysages Système d'information sur l'eau en France ; Water Information System France
ICELAND	Interactive map solution – Icelandic nature
ITALY	GELSO – best practices data base for the local sustainability
NORWAY	Norway digital – Norwegian Spatial Data Infrastructure The Artsdatabanken biodiversity information resources
SWEDEN	WaterInformationSystem Sweden (WISS)
EEA (EU LEVEL)	Near real-time air quality system (Ozone web and PM web)

4. NEXT PROJECT STEPS

Following the above detailed activities, the next phases of the project are going to develop the following actions:

- an assessment of the State of Play at National level
- continuation of the work addressed at identifying and collecting Best Practices. We remark that though being the starting point of the project, both the State of Play at National level and the Best Practice Catalogue are considered as a "living" tasks till the end of the project and effort will continue towards them
- the synthesis of the State of Play and its validation at European level, from which to derive lessons
- the assessment of the structure of the project deliverables "ICT roadmap" and "Guidelines on ICT supporting environmental and monitoring" and a progress in their implementation
- the development of the network at National and European level, with actions such as the organisation of meetings and workshops and the active participation in the main events related to the mission of the project. Particular effort will be dedicated to the country scope enlargement of the network, involving stakeholders representatives of countries not yet included in NESIS
- linking other initiatives and actions related to the implementation of SEIS, such as the identification of priorities for further ICT initiatives of DG INFSO and the establishment of a proactive collaboration with the Technical Group of the SEIS Task Force.

REFERENCES

Crompvoets J., Rajabifard A., van Loenen B., Delgado Fernandez T., (Editors). A Multi-View Framework to Assess Spatial Data Infrastructures. Digital Print Centre, The University of Melbourne, pp: 403, 2008.

Vandenbroucke D., Janssen K. Spatial Data Infrastructures in Europe: State of Play 2007. Summary report by the Spatial Applications Division, K.U. Leuven R&D. 2008.

European Commission. Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions - Towards a Shared Environmental Information System (SEIS). COM/2008/0046 final. Online: eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52008DC0046:EN:NOT, 2008.

Approach to Build a Soil Information Portal for Europe Based on the PortalU® Technology

S. Uhrich[a], **M. Klenke**[a], **F. Kruse**[a], **C. Giffei**[a]

[a] *Coordination Centre PortalU at the Lower Saxony Ministry of Environment and Climate Protection, Archivstr. 2, 30169 Hannover, Germany (kst@portalu.de)*

Abstract: The access to environmental information has become an important concern in the recent years for both public and private bodies within Europe. Especially the web-based supply of the huge amount of spatial environmental data deserves particularly attention because high organisational efforts and financial expenses are necessary to improve the access to this kind of data. While the INSPIRE Directive and its Implementing Rules (IR) (European Union 2007) give the framework to establish a European spatial data infrastructure, vital obstacles in reference to harmonization and interoperability of data and services as well as in reference to the organisational structure are not removed yet. The project GS Soil aims to make a contribution to remove these obstacles by establishing a European web portal for soil information (GS Soil Portal). The main objectives of the project are to build up the base for an easier reuse of information about soils by improving interoperability of spatial data sets and services. Especially the identification of issues and the implementation of solutions for reducing the barriers to interoperability of spatial data sets and services are focused. GS Soil concentrates thereby on soil and soil related themes enumerated in annexes I to II of the INSPIRE Directive (European Union 2007).

Keywords: Environmental information; Soil data; Spatial data infrastructure.

1. INTRODUCTION

The availability and accessibility of environmental information has become a key concern for public and private bodies within Europe in the recent years. The European Directive 2003/4/EC (European Union 2007) on public access to environmental information, the INSPIRE Directive 2007/2/EC for establishing an Infrastructure for Spatial Information (European Union 2007) as well as the newest initiatives of the European Commission like the Shared Environmental Information System (SEIS) (COM 2008) and the Single Information Space for the Environment (SISE) (O'Flahery 2008)) emphasizes the European-wide need to improve the access to environmental information. Especially the web-based supply of the huge amount of spatial environmental data deserves particularly attention because high organisational efforts and financial expenses are necessary to improve the access to this kind of data. While the INSPIRE Directive (European Union 2007) and its Implementing Rules (IR) give the framework to establish a European spatial data infrastructure, vital obstacles in reference to harmonization and interoperability of data and services as well as in reference to the organisational structure are not removed yet.

The eContentplus project GS Soil is currently under negotiation and will likely start in May 2009. EContentplus is a multiannual Community programme from the European Commission DG Information Society and Media with the objective to make digital content in Europe more accessible, usable and exploitable (DG Information Society and Media 2008). GS Soil is thereby be allocated to the area of geographic information, where the focus is set on the aggregation of existing national datasets into cross border datasets, which will serve to underpin new information services and products, in particular with a view to reducing barriers related to one or more of the specific themes mentioned in annexes I-III of the INSPIRE Directive (European Union 2007). The focus of GS Soil is thereby set on soil and soil related data. This paper describes the project structure and the technical base to build up the European GS Soil Portal.

2. THE PROJECT GS SOIL

2.1 Introduction

The project GS Soil aims to make a contribution to remove obstacles referring to harmonization and interoperability of soil related data and services as well as in reference to the organisational structure by establishing a European network to improve the access to spatial soil data for public bodies, private companies and citizens. The project will focus on data organisation, data harmonisation as well as semantic and technical interoperability with the objective to produce seamless spatial information in terms of INSPIRE (European Union 2007). Both the description and harmonisation of European spatial soil data and the the operation of a corresponding spatial data infrastructure will take centre stage. Within the project 34 partners from 18 European Member States are involved. Soil data are thereby provided for all 18 states mainly on national level and partly on regional level. That means 67 % of the 27 European Member States will provide soil data for the project. These data build a sufficient base to analyse and improve the access to the different kinds of digital content.

2.2 Project structure

In order to achieve the aims of the project, GS Soil is structured in seven work packages (WPs): WP1 for project coordination and networking, WP2 for building up the content provision framework, WP3 with focus on metadata, WP4 with focus on harmonisation and semantic interoperability, WP5 for establishing of an integrated network and a soil portal, WP6 for evaluation and sustainability and WP7 with focus on dissemination, awareness and clustering activities with other projects. WP2 to 4 will set up on each other. The soil and soil related data will be analysed, necessary metadata will be identified and provided and specific datasets will be systematically harmonised. The improved access to soil information by the GS Soil Portal is the main objective of WP5, while WP6 focuses on user needs and the long-term perspective of the GS Soil Portal.

At the beginning of the project specific and generic requirements for soil information, services and products have to be identified by a range of user communities and stakeholders. This requirement analysis will result in a soil inventory and theme catalogue, which documents the current state and ability of data providers to meet the goal of data harmonization. Based on this, specific content requirements and characteristics will be defined as fundament for the establishment of harmonized soil data sets.

As a fundament for harmonizing different kinds of soil data, basic information about the data sets have to be made available in the form of methodical metadata, which describe the underlying data models in depth, and which comply with a generic framework for developing application schemata for the relevant data. Therefore a schema for describing spatially linked soil data and database services to comply with the INSPIRE Directive will be developed, which meet the needs of data users for a harmonised, interoperable EU-wide, national and regional soil data infrastructure. A soil oriented metadata structure profile will be developed following the INSPIRE IR for metadata, other international, and national standards, like the ISO 19115:2003, and the needs of the data users.

In the context of data harmonization a framework will be developed that enables existing soil datasets to be linked up from one country to another. Therefore data specifications will be developed and made available as basis for designing and testing of transformations or translations. The level of spatial data consistency depends to a large proportion on user level harmonization efforts. Such efforts will be thoroughly analysed in order to identify, to which degree pre-harmonization is needed and how it can be implemented. However, particular focus will be given to the development of user level transformation services in the context of the INSPIRE harmonization framework. Exemplary services will be developed with the objective to present comparable and meaningful data portrayal. The focus on soil map legends and soil inventory data is important because attribute and property data are crucial for developing evaluation and transformation services. The results will enter in a best practice guideline for soil data specification development under INSPIRE.

2.3 The GS Soil Portal

InGrid®, the technology of the German Environmental Information Portal PortalU® (Voegele at al. 2007), will be used as technical base in the project to build up a European GS Soil Portal. The European web portal, GS Soil Portal, is used to bundle the decentralized distributed soil information of the 18 states (Fig. 1).

In the GS Soil Portal all soil related information, web pages, databases as well as data catalogues, can be searched by simple or advanced search queries according to the requirements of the user. The results of a query are displayed in a ranked result list independent from the data source.

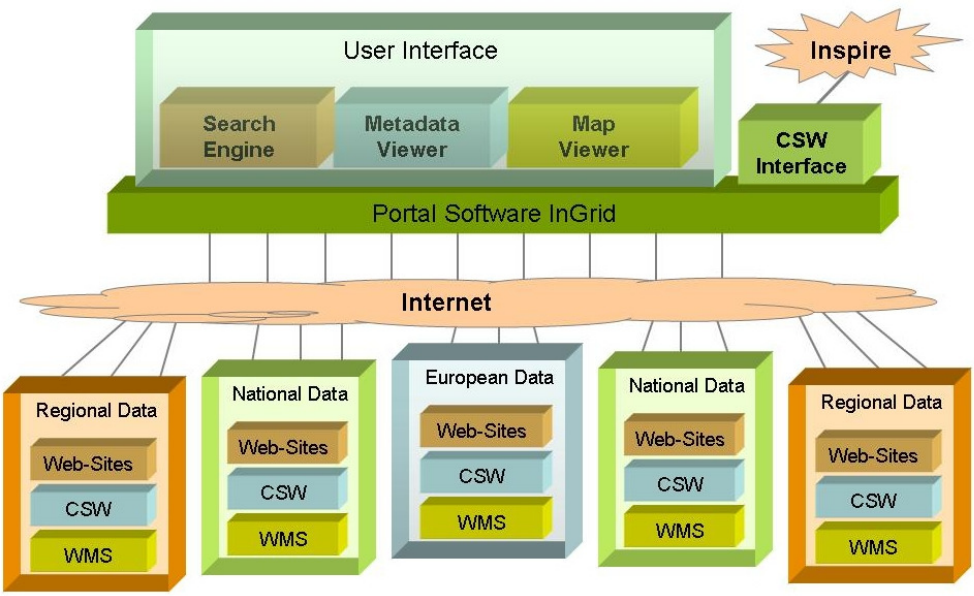


Figure 1. Architecture of the PortalU® technology InGrid®

The technology provides different kinds of interfaces to data sources (iplugs) and also interfaces for the transfer of information to other systems like INSPIRE. The most important iplug / interfaces in reference to the GS Soil project is the CSW-2.0-interface (Fig. 1). The OGC conform CSW-2.0-interface is based on the ISO standards 19115, 19119 and 19139 and the IR metadata of INSPIRE. This interface makes the exchange of spatial metadata in both directions possible. With the CSW-iplug spatial metadata sources (data catalogues) can be connected to the GS Soil Portal, while on the other hand the CSW-interface can be used to transfer spatial metadata from the GS Soil Portal to external systems. In reference to embed the GS Soil in a major spatial data infrastructure like INSPIRE the standardised CSW interface and further standardised interfaces play an important role.

The PortalU® technology provides a hierarchical structure of partners and data providers in order to consider the federal administrative structure in Germany. According to this, the federal government and the states (Bundeslaender) act as partner in PortalU®, the single data providers are accordingly subordinated to the federal government or the referring state. This structure can also be easily used on European level for GS Soil: each state will act as partner and the data providers are allocated to the referring country. In the GS Soil Portal all soil related information will be made available and accessible. Spatial soil data from OGC compatible Web Mapping Services (WMS) and Web Feature Services (WFS) will be visualized in the map viewer.

The technology has a modular structure and is thereby easily extendible for the specific needs within the project. It is based on open source components and internal developments. Therefore it can be used without external licence costs in the project. For all tasks within the project the GS Soil web Portal will be used as platform for an improved access to the soil data.

REFERENCES

COM (2008): Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions – Towards a Shared Environmental Information System (SEIS), SEC 2008 111, SEC 2008 112, 0046 final.

DG Information Society and Media (2008): eContentplus programme, *http://ec.europa.eu/information_society/activities/econtentplus/programme/index_en.htm*, 9.2.2009

European Union (2003): Directive 2003/4/EC of the European Parliament and of the Council of 28 January 2003 on public access to environmental information and repealing Council Directive 90/313/EEC – in: Official Journal of the European Union L 41, 26.

European Union (2007): Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) – in: Official Journal of the European Union L 108, 50.

O'Flahery, J. J. (ed.) (2008): Towards a Single Information Space in Europe for the Environmnet, Expers Consultation Workshop Brussels 15.2.2008, organised by ICT for Sustainable Growth Unit of the EU DG Information Society and Media, *ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/sustainable-growth/sise-workshop-report-08_en.pdf*, 9.2.2009

Voegele, T., M. Klenke, F. Kruse, H. Lehmann & C. Giffei (2007): PoralU® - a Tool to Support the Implementation of the European Environmental Information Directive, 21st International Conference on Informatics for Environmental Protection, 2, EnviroInfo Warsaw 2007, Aachen.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Application of SEIS and INSPIRE in Slovenia

J. Urbančič [a], **A. Veršič** [b], **U. Kušar** [c]

[a] Environmental Agency of the Republic of Slovenia, Vojkova 1b, 1000 Ljubljana, Slovenia (jelko.urbancic@gov.si)

[b] *Environmental Agency of the Republic of Slovenia, Vojkova 1b, 1000 Ljubljana, Slovenia (aversic@gov.si)*

[c] *Environmental Agency of the Republic of Slovenia, Vojkova 1b, 1000 Ljubljana, Slovenia (urska.kusar@gov.si)*

Abstract: The implementation of the environmental information system in Slovenia is based on the idea of Single European Information Space. Shared Environmental Information System (SEIS), GMES and INSPIRE activities are part of this initiative. Ensuring high-quality environmental data for all target groups is an important item in the mission of the Environmental Agency of the Republic of Slovenia. It is being carried out through four web applications. The first one is used to ensure relevant information based on environmental indicators and the other three (Environmental Atlas, the metadata portal and the application for downloading spatial data) are used for geographical presentation of all other available environmental data.

Keywords: Environmental data; Environmental information; Metadata; Environmental indicator; Geographical information system; GIS;

1. INTRODUCTION

The directive INSPIRE was adopted in February 2007. Pursuant to the directive, European Union Member States will have to set up an infrastructure for spatial information to support the EU's environmental policies, as well as policies and actions which may have a direct or indirect impact on the environment. The infrastructure will have to include metadata, spatial data sets and spatial data services ranging from the local to global level in a manner enabling multi-purpose utilisation.

Side by side with the above mentioned directive goes a trend towards enabling access to data at the global level through Web services for spatial data. In most European countries, a number of distributors of spatial data facilitate the accessibility of such data through the Web Map Service (WMS) and Web Feature Service (WFS). In 2002, the Survey and Mapping Authority of the Republic of Slovenia started with systematic set-up of a distribution environment for the survey and mapping service, as well as its integration into the Slovenian e-government system, which enables direct computer access to survey and mapping data to different user groups, such as government and local authorities, particular users (survey and mapping companies authorised to carry out survey and mapping services, notaries, lawyers, real estate agencies, building managers, etc.), other companies and natural persons [Zbornik, 2006].

The described system is designed for state administration; currently, other users may only examine data and it is only the first step towards Web services enabling broader use. The aim of WMS and WFS services is to provide individual users with an updated source of information that can be combined with data from other sources, thus enabling the real time creation of information needed by the user.

Among the mission and strategic objectives of the Environmental Agency of the Republic of Slovenia, National and international legislation and conventions e.g. Aarhus Convention, OECD's Council Recommendations, EU Directives and Communications the following are stressed:

- providing quality environmental information, analysis and basis for decision making

- ensuring greater public safety, and efficient methods to inform the public about environmental issues and environmentally hazardous situations
- introducing simple, citizen-friendly administrative procedures
- performing regular work with efficient operations and the lowest possible costs

Environmental indicators are among the most applicable tools used for the purposes of environmental reporting. Based upon numerical data demonstrating the status, specific characteristic or development of a certain phenomenon, they can warn of specific issues. They help us to measure and determine the quantity of diverse data constituting a complete data collection. The indicators are, in fact, data that have been collected and presented in an agreed manner, with the purpose of establishing the connection between the existent data and the targets of the environmental policy. Appropriately selected indicators that are based upon an adequately extended time series of data can provide a demonstration of key environmental trends [Rejec Brancelj et al., 2003].

On other hand is the Agency providing access to quality environmental data through WMS and WFS services and thus follows its mission and strategic objectives. Up to date and accurate information, as well as the simplest possible access to needed data, are important to every user; therefore the application of Web-based services is perfect way of transmitting spatial data.

The paper describes the back office and front office processes, the solution of public environmental system based on international and internal standards to enable an automated information flow that can be used for internal scientific communication and also for providing relevant information to the public. The presented paper describes two examples of web based environmental information system, one is providing general environmental information based on environmental indicators, and the other is presenting geographical based information incorporating the above mentioned principles.

2. ENVIRONMENTAL INDICATORS IN SLOVENIA

The "Environmental Indicators in Slovenia" is a system of indicators published by the Environmental Agency of the Republic of Slovenia as a web site (http://kazalci.arso.gov.si) and occasionally as a printed report [Kušar et al., 2006]. The system is organized and financed as part of national reporting on the state of the environment in accordance with obligation from Article 106 of the Environment Protection Act. They also comply with the public's right to be informed on the state of the environment and the efficiency of environmental policies. Because of the standardized structure (Figure 1) indicators are also important collection of information on available data sources and data flows related the environment.

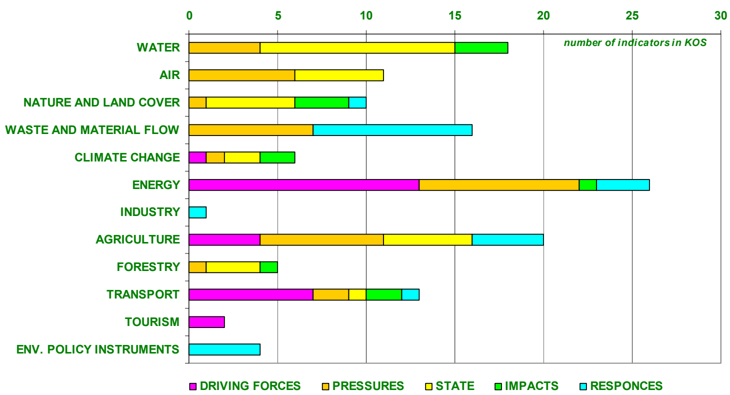


Figure 1: Number of indicators by theme in the Environmental Indicators in Slovenia system

At the moment 132 indicators are published as a part of Environmental Indicators in Slovenia system. For the purposes of a more transparent overview, the indicators are grouped in twelve thematic clusters – sections. The latter treat environmental media (e.g. water, air, etc.), environmental issues (e.g. depletion of the ozone layer and climatic changes, protection of nature and loss of biodiversity, waste generation and management) and integrated indicators under a system for the formation of sector policies (e.g. indicators relating to transport, agriculture, tourism, energy and instruments of environmental policies).

2.1. The structure and the presentation of an indicator

During the development of the environmental indicators system much importance was given to the development of the structure of an indicator. Based on the European Environment Agency's model [2003] we elaborated the indicator template with easy-to-understand guidelines on the form and content of each indicator's section [Kušar et al., 2006].

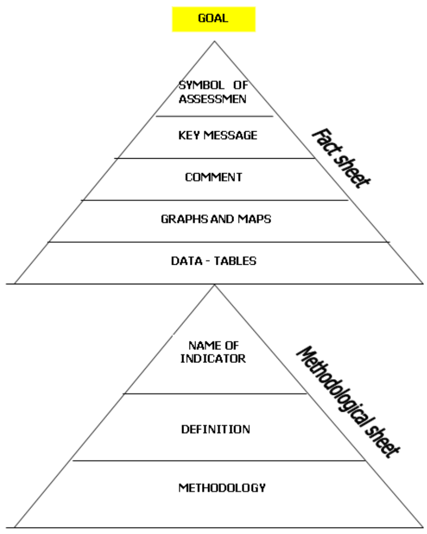


Figure 2. Model of an indicator

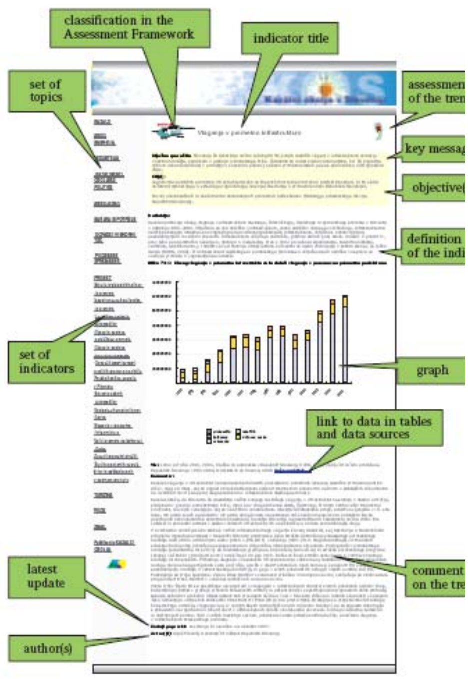


Figure 3. Indicator's presentation

Each indicator is determined by a **definition** providing basic information on the methodology of conducted measurements and the manner of demonstrating the indicator in question. The indicators rely on internationally verified methodologies and are thus, as a rule, internationally comparable. In their preparation we have mostly used methodological sheets for indicators as drafted by the European Environmental Agency [2003]. Where so required by a certain phenomenon and the method of its monitoring, the accessibility of the data or any other technical factor, the EEA methodology has been adapted to conditions specific to Slovenia.

In evaluating the development of a certain phenomenon it is of crucial importance to be aware of an envisaged trend and intensity of the development. Therefore each indicator is accompanied by a **goal**. As a rule, the required trends are taken from the fundamental programming document in the field of environmental protection, i.e. the National Environmental Protection Programme 2005-2012 as well as other sector strategies and programmes. Also taken into account are the strategic targets set by the EU and indicated in the relevant Directives, the EU's Environmental Action Plan and the objectives that Slovenia has committed to by signing international treaties.

The quantitative values of a given indicator are expressed mainly in annual values and shown with **figures and maps**. The latter are accompanied by **explanations** interpreting

the development and likely reasons, as well as implemented and envisaged measures for enhancing or preserving the current state of the environment. Each explanation is further accompanied by **a symbol** providing an **assessment** and **key message** of individual indicators. Each assessment is prepared by an author for each individual indicator and represents a "summary" of expert assessment of a trend with respect to presented data and set targets.
Contributing to the transparency of the applied methods of monitoring the selected indicators is the Section entitled **Data and sources** or **Methodology** accompanying each indicator. Special attention is given to use the exact reference of the database used, the institution and timing of data capture. This allows us to establish direct links with other web tools which provide data (e.g. on-line data) or metadata from certain data source e.g. Catalogue on Data Sources on the Environment (http://kpv.arso.gov.si/). In addition to the **tables containing the data**, this section also offers a more detailed description of the data sources used and provides further methodological notes.
Links to the connected indicators section enable users to explore the connection between indicators, primarily following the D-P-S-I-R assessment framework.

2.2 Background process

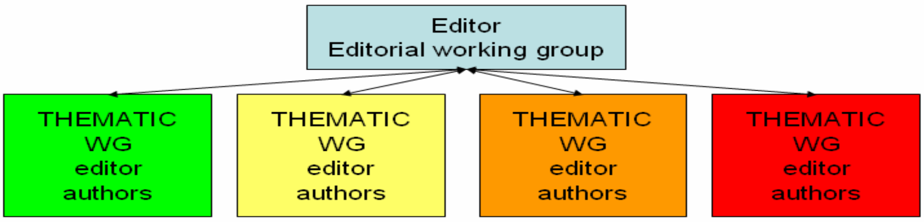


Figure 4. Environmental indicators in Slovenia system's organisational scheme

Environmental indicators, as used in the Environmental indicators in Slovenia system, are composed from data and expert assessment. Data used are highly aggregated (usually annual sums or averages) and abstracted from aproximately 50 different databases and other data sources. They are updated annually or bi-annually. To assure consistency and timeliness of presented information, efforts are put into development suitable organisational structure, workflow procedures and standards.

The basis for the composition of the indicator set (selection of indicators) is the assessment framework which helps to define the functions of respective indicators. The used five-partite assessment framework, so-called DPSIR Assessment Framework includes the following set of concepts: Driving forces – Pressures – State – Impact – Responses, where each individual set conveys its own meaning. The selection is done by consultation process in which editor, thematic editors and thematic working group is involved.
The responsibility for each indicator's content is primarily on the indicator's author. Authors are nominated by

Figure 5. Indicator's graphs and tables editing web form

indicator's system editor and thematic editor. Issues of an indicator selection, methodology and presentation are discussed in thematic working groups. Authors of indicators are typically experts covering thematic issue of the indicator within institutions responsible for monitoring selected thematic, which usually includes management of relevant databases.
Usually once a year in the agreed period, which depends on the nature of data collection, the editor sends a request for update of the indicators to the relevant authors, together with pre-prepared template with detailed and easy-to-understand instructions on the form of indicator for both, text and data part of the indicator.
As web application [KOS] enables interactive presentation of graphs and tables, the data must be prepared following table – template (MS Excel file) from which a macro creates xml file, which is inserted into web form (Figure 5) and through Zope application, the graphs and tables are created and presented on the web site. (Figure 3).

3. ACCESS TO THE SPATIAL DATA

3.1 Metadata Portal

There are two approaches. One is to view or read the data and second is to access or download the data. To read about the data users can use Metadata Portal where are all information about the data which can be downloaded or retrieved using WFS.

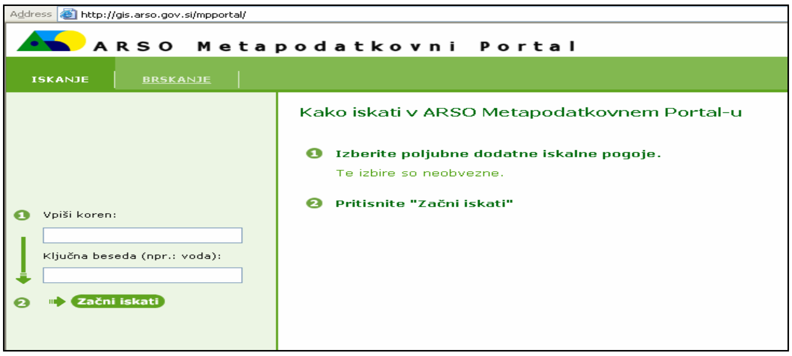


Figure 6: Metadata portal

Metadata portal enable to search or browse through the portal. Users can search by keyword or as a root. If we want to browse, the metadata are sorted by different categories.

Figure 7: Example of metadata description and functionalities

When we choose the data there is a link to more detailed description, link to Environmental Atlas, to WFS or to download the data using GISRepSystem.

3.2 Environmental Atlas

The Environmental Atlas (EA) is a Web service which enables users to examine spatial information via the Internet. Through this service, the Environmental Agency of the Republic of Slovenia provides access to relevant environmental content to a wide range of users.

Using geographic display and spatial queries, user can obtain answers to questions on the position of a certain phenomenon in space, its location and its relation to other elements.
The location queries and the cross section of layers provide data, including position coordinate, plot number, possible location in ecologically significant or protected areas, characteristics of the relevant area, possible vicinity of water resources and hydro-geological characteristics. The display with marked location can be printed or, together with observations or other remarks on a certain location, sent by e-mail to co-workers, customers, authorised persons, etc. The EA tools enable the user to create and edit layers, and to import/export layers in the .gml and .kml format.

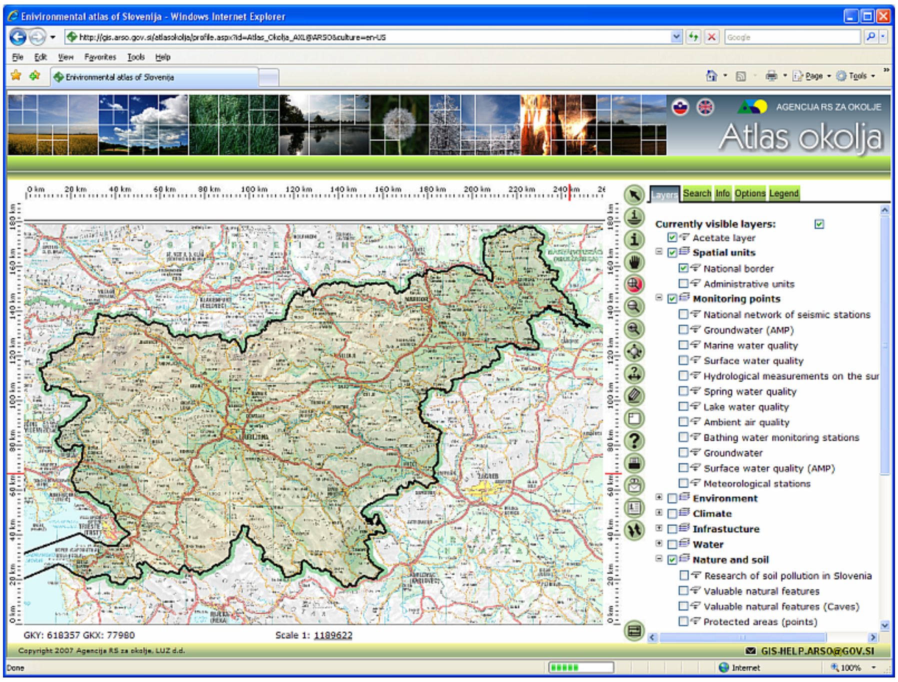


Figure 8: Environmental Atlas

3.3 WFS – Web Feature Service

WFS enable access and modification of data using HTTP protocol. The enabled operations are comprise, insert, modify, delete, lock, query and search. [Okoljski GIS]

The Agency's "GISRepSystem" solution enables simple publication of data from the Oracle Spatial format, using the WFS standard. It enables installation on any J2EE-compatible application server. Agency's data are stored in Oracle Spatial format, while the solution provides for the creation of flexible data sources.

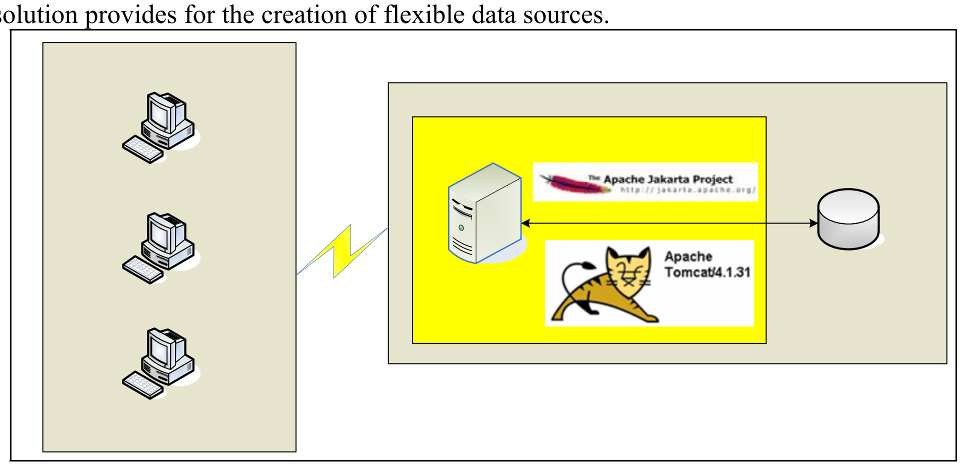


Figure 9: Installation of the GISRepSystem in the information system

Through the GISRepSystem, users can obtain data via the WFS service. Access to the data is available through the Java application WFSCLientUI enabling the transfer of data in .shp or .gml format. In addition some advanced functionality include:

- selection of one or more WFS data sources
- limitation of copied data according to selected columns and given criteria
- frequency of updating data and output format
- merging data sources into larger groups, called profiles
- background functioning of the program, which facilitates uninterrupted daily performance of replications

The second option through WFS protocol enables connection of user application to WFS service using its URL. In this way, the user obtains data in real time and directly from the distribution server.

4. CONCLUSIONS

The concept of Shared Environmental Information System is feasible. The presented examples are operational applications part of information system of Environmental Agency of the republic of Slovenia. The key success factor of working information system is a well defined protocol. In the case of global sharing information the exchange protocol must be based on international standards, while in other cases internal or community standards are sufficient. All mentioned above is well known from information system theory and all three presented geographic system applications are good practice examples of efficient data exchange and enabling also basic on line data analysis. The environmental indicators portal is a good practice of systematically presented “soft” environmental information. It can be used as well for providing public information and it can be also used in reporting on state and trends of environment.

REFERENCES

Ažman, I., Vzpostavitev slovenske prostorske podatkovne infrastrukture. Magistrska naloga. Ekonomska fakulteta, Univerze v Ljubljani, Ljubljana, 2005

Environmental Atlas, <http://gis.arso.gov.si/atlasokolja>

European Environment Agency, Environmental signals 2002: benchmarking the millennium: European Environment Agency regular indicator report, 147 pp., 2002.

European Environment Agency, EEA core set of indicators, revised version April 2003. 2003

INSPIRE, Directive 2007/2/EC establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), 2007/03/14

KOS, Environmental indicators in Slovenia, <http://kazalci.arso.gov.si/ >

Kušar U., Kovač N. and I. Rejec Brancelj, Uvod. Kazalci okolja 2005. Ljubljana : Ministrstvo za okolje, prostor in energijo, Agencija Republike Slovenije za okolje, 6-10, 2006.

Metadata Portal, <http://gis.arso.gov.si/mpportal>

Okoljski GIS, <http://gis.arso.gov.si/>

Povše U. and I. Rejec Brancelj, Izkušnje in prihodnost poročanja o okolju, Vključevanje in povezovanje statistike v informacijsko družbo = Integration of statistics in the information society : proceedings, 369-379, 2002.

Rejec Brancelj, I. and U. Povše, Kazalci kot orodje za spremljanje okoljskega razvoja, Statistika kot orodje in vir za kreiranje znanja uporabnikov: zbornik = Statistics as a tool and source for users' knowledge creation: proceedings, 465-473, 2003.

Zbornik, 2006: http://www.zrc-sazu.si/slogis/zbornik.htm (24. 6. 2006)

Session 3

ICT for monitoring and controlling of energy efficiency and security, modelling of air pollution, climate changes to support decision making in environment protection

Organized by **Emil Pelikán**, Vladimír Bízek, Giorgio Guariso and Ivan Obrusník

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pilmmann, I.Holoubek, R. Legat (eds.)
Masaryk University, 2009

A New Architecture for Reduction of Energy Consumption of Home Appliances

A. Capone[a], M. Barros [b], H. Hrasnica[b], S. Tompros[c],
[a] Politecnico di Milano, Piazza L. Da Vinci 32, Milan, Italy
antonio.capone@polimi.it
[b] Eurescom GmbH, Wieblinger Weg 19/4, Heidelberg, D-69123, Germany
{barros,hrasnica}@eurescom.eu
[c] Keletron LTD, 49 D. Glynou, Thessaloniki,GR-54352, Greece
stombros@keletron.com

Abstract: This paper presents the recent results achieved by AIM consortium [AIM, 2008] in developing and demonstrating a new information and communication technologies (ICT) architecture for modelling, virtualising and managing the energy consumption of home appliances. The architecture aims at fostering a harmonised technological frame for profiling and optimizing the energy consumption patterns of home appliances. The intention is to offer users a number of standalone- and operator-based residential services that will allow them to manage efficiently the energy consumed in households. To make these services possible the frame adopts a generalised method for household appliances management, which is based on an accurate modelling of operational modes of appliances and the ability of the home network to switch on or off some of their internal functions without limiting their control just to the active or stand-by states. As a pilot application, the appliances being considered for the first implementation of the architecture are: white goods, e.g. refrigerators, ovens, washing machines, dryers; audiovisual equipment, e.g. TVs, DVDs, Set-top-Boxes; and home communication devices, e.g. wireless routers, DECT phones, residential gateways, modems. The requirements defined for the architecture concern mainly usability aspects of power management functions, integration with the home network and service deployment. The final result provided by the project is a system with enhanced home network architecture, incorporating services for home appliances energy consumption monitoring and management and using a generic technology. The two major challenges are energy saving and the architecture's long-term sustainability.

Keywords: Energy efficiency; Residential gateway; Energy management; Energy consumption's reduction.

1. INTRODUCTION

The main concept of the architecture developed in AIM project [AIM, 2008] is to offer a harmonised technology for managing in real time the energy consumption of appliances at home, interworking this information with communication devices over the home network to make it available to users through home communication networks in the form of standalone or network operator services. By using ICT to achieve technology-driven energy efficiency gains, AIM is supporting the European Union's Action Plan on Energy Efficiency [EC, 2006]. Figure 1 illustrates the conceptual model that serves the logical basis to AIM architecture. The main innovation in managing the energy of household appliances is the bridge between home communication and power distribution networks with the aim to control the power distribution through communication services.

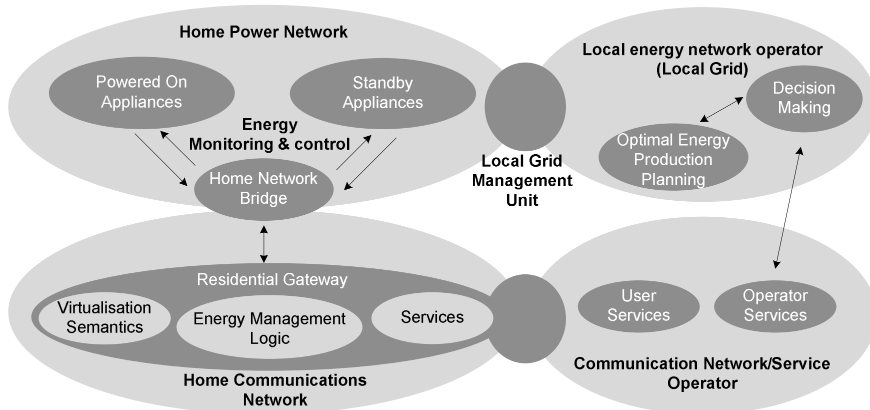


Figure 1. Conceptual model of AIM Architecture

Network operators may use the interfaces of the residential gateway to implement services for mobile and fixed terminals featuring remote energy monitoring and control of the home environment.

Power distribution network operators have particular interest to monitor the energy consumed by large blocks of users on macroscopic level. Accessing households through such a system is an efficient and cost effective way of accomplishing such task.

Residential users may control their environment through the service interface of the gateway that is able to get connected with any type of home terminal, like e.g. wireless PDA, embedded devices, et.. Moreover, the system is able to collect additional information from the environment through a sensor network and create user profiles in order to perform a partially automatic configuration of the energy management policies. Home terminals distribute commands to the appropriate appliance via the EMD, affecting its energy consumption attributes.

The paper is organized as follows. Section 2 describes the AIM general architecture. Section 3 details some usage scenarios and energy saving policies. In section 4, the user interaction with the system are explained and discussed. The conclusions are presented in section 5.

2. AIM ARCHITECTURE

AIM architecture is presented in Figure 2. AIM bridges the outdoor and indoor networks with the view to provide the means for controlling the functions of the household appliances through a number of different applications addressing three user categories: i) Residential users; ii) Network operators; iii) Energy generation utilities.

The indoor (home) network is bridged with the outdoor networks through the "AIM system logic" (see Figure 3) whereby users are enabled to manage the functions of household appliances and control the energy consumed in their households.

The AIM system logic is the main building block of the architecture, interconnecting the home network, the outdoor networks and the software substrate for the implementation of energy saving applications. The AIM gateway appears as a building block of the AIM system logic, for what it may optionally host part or the whole of the AIM system logic, or being used as a passive component while the service logic is hosted on the operator service platform. The AIM gateway selects and conduits information to the proper device interface, applies the necessary centralized control logic and enforces rigorous communication encryption. The apparatus that constitutes the local hub of the energy control system is the Energy Management Device (EMD), which is an independent functional entity that conveys control logic for both active and stand-by appliances and energy management functions integrated through a multimode of communication interfaces with the home network and the

AIM gateway. The EMD is controlled by the gateway, using a bus interface that grants access to multiple EMDs from a single access-point, either locally or remotely via an operator network.

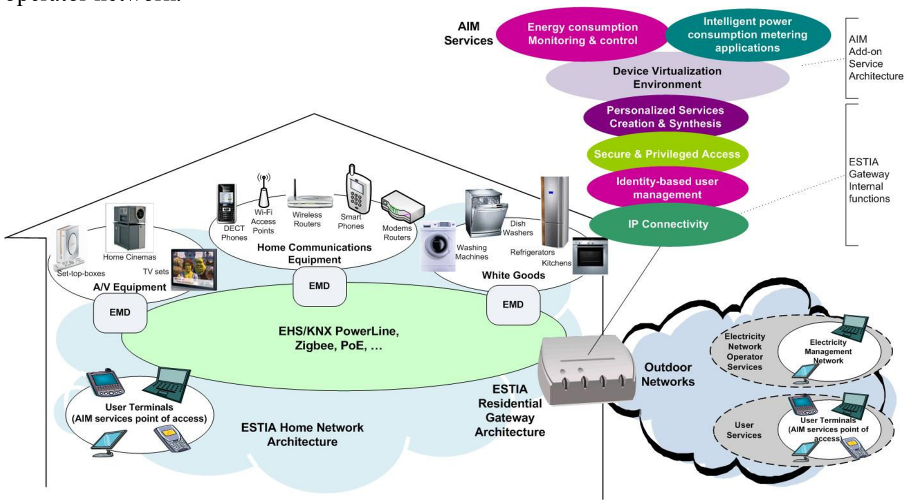


Figure 2. AIM Reference Architecture

2.1 Energy Management Device (EMD)

The EMD must have a unified architecture, which will feature generic interfaces towards the household appliances, the power network and the home network. Due to its generic architecture it can be implemented as a standalone external device, integrated in the AIM gateway, as well as an internal module of the appliance. The EMD shall be accessible either locally, through the AIM gateway or via external operator networks. Privacy and confidentiality of user data circulated in the outdoor networks will be ensured by the application of proper encryption of messages exchanged between the EMDs and the AIM network.

The EMD will offer three generic-purpose interfaces: one towards home communications networks, one towards the mains power network and one for connecting to internal digital control buses of household appliances. With these interfaces the system will be able to integrate with virtually any network environment or household appliance and will provide two types of power management logic:

- Power monitoring, or power metering functions that are applied to power electronics of the household appliances, an encoding logic that turns measurement results into digital values and a monitoring logic that buffers the obtained measurements following user configuration commands.
- Power control, or control logic for selecting which of the several external interfaces will communicate to the household appliance, taking into account the user commands as they have been decoded and submitted by the enforcement logic of a given appliance.

2.2 AIM Gateway

AIM adopted ESTIA [ESTIA, 2006] gateway to use as AIM gateway for scalability, upgradeability and openness reasons, because it is based on the open services execution framework of OSGi. AIM gateway is composed by 3 modules:

- Machine-to-machine interfaces module, which delivers a unified methodology and a common API for the implementation of gateway-based services, incorporating the connected appliances. It defines a novel mechanism that consolidates the different access and communication technologies under a single umbrella.

- Identity management module: responsible for user authentication/identification and for providing personalised applications to the user.

- Services synthesis module: allows the creation of new composite services based on existing service primitives, which are provided according to the user profiles, device profiles and the associated policies.

Most of the energy management functionality will be hosted by the gateway, including: appliance capabilities discovery; appliances and user profiling; virtualisation environment that enables the residential user to access household's energy resources and exploit them in defining energy management processes; management of the user interface; make the energy consumption statistical data towards the outdoor networks anonymous; energy monitoring and management, through providing to the user services with APIs for communicating with the EMD, the device that mainly performs energy monitoring and management; communication between indoor and outdoor components harmonisation.

2.3 Logical Interfaces

Residential and outdoor users as well as third parties shall have access to AIM services through applications compatible with any user terminal type, e.g. wireless/wired terminals such as mobile phones, PDAs and PC based consoles. The user interface coincides with the user application hosted on the user terminal and allowing users to access the services of the home network.

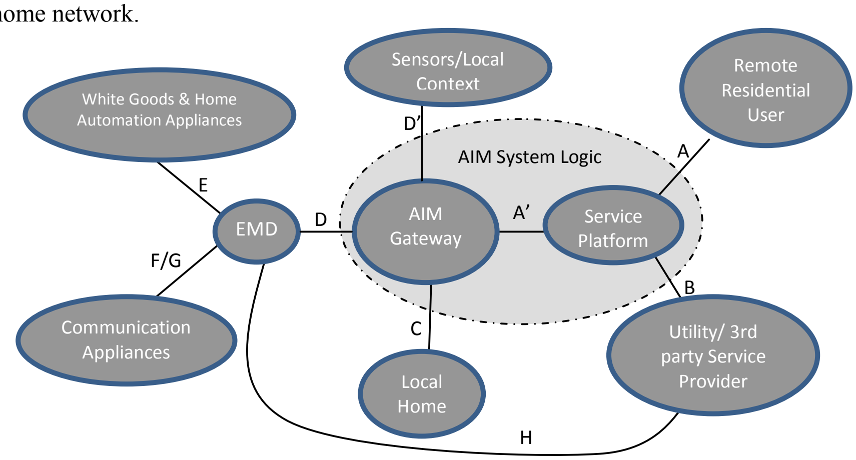


Figure 3. Logical Interfaces of AIM Architecture

Figure 3 presents the logical interfaces of AIM architecture, which are communication links that bind the components altogether. The term 'logical' comes from the logic that conveys their operation, which depends on the type of interconnected components:

- Interface A is a remote access http web browser-based interface that allows the users to access the AIM system for both control and monitoring when they are outside the home network. It should give access to all the regular control and monitoring functionalities, with possibility to restrict access to some of these functionalities only to specific categories of users. Interface A' allows to transparently traverse the main home network gateway whilst maintaining security, because it avoids to open a permanent backdoor that could be taken advantage of by potential intruders.

- Interface B is a logical interface and is independent from the way how the information is transported between both entities that interconnects. Its main function is to provide the required information between the utility and the household.

- Interface C is the interface between the local users and the AIM gateway, representing the logical connectivity of the residential user with the services of the home network and conveying information that allows the user to control energy consumption of the home environment.

- Interface D is the interface between the EMDs and their controlling gateway. Since the AIM domestic network is essentially hierarchical, all flow control will be governed by an AIM gateway that will be the central coordination point and can accept commands from external actors and send them to a specific or a group of EMDs in the system.

- Interface E is the interface between the EMD and a white good. The EMD communicates with AIM gateway over a standard AIM protocol and translate the information from/to the AIM gateway into the proprietary protocol of the white good.

- Interface F/G is the interface between the EMD and communication appliances, e.g. phones, routers, or audio/video equipment, e.g. radio and TV sets. To have power measurements and saving features sometimes additional devices will be needed. have multiple interfaces to communicate among themselves and between other devices and use. The proprietary commands of the communication devices will be translated into AIM commands by the EMD. The data exchange between the EMD and the AIM gateway will use AIM protocol.

- Interface H is an option to connect the EMD directly to the service providers for when there is no possibility to have a gateway with interface for remote control installed at the home network.

3. USAGE SCENARIOS AND ENERGY SAVING POLICIES

The functionality that will be offered by the AIM system is intended to be used by three different types of users, namely utility providers, telecom operators and the local users themselves. In the following we give some narrative examples of the usage scenarios and possible energy saving polices that could be enabled by the system.

3.1 Local users

John is a typical worker that goes to work quite early in the morning and comes back home in late evening for dinner and the night his main goal when managing the system is to minimize incurred costs. John wakes up at 7.00 a.m. and, since we are in winter, would like a pleasant temperature in the rooms where is going to live in while he prepares breakfast and dress himself before leaving. The system increases the temperature from night value of 18 °C to 22 °C in the bedroom, bathroom and in the kitchen, while it keeps the temperature low in the living room because John does not use this room in the morning. The system also activates the appliances that John likely need to use in the kitchen and bathroom, keeps in stand-by mode those that may be used, while it leaves off all the others.

Before leaving, John schedule some activities that the system has to execute before he is back, like cooking the diner, run the washing machine, operate the vacuum cleaner robot, etc. When John is not at home the temperature is reduced and the activities that he scheduled are executed trying to minimize energy cost based on real-time price information provided by the utility. The system can take autonomous decisions considering the cost minimization goal and the quality constraints provided by John. For example, diner needs to be cooked and also warmed for 7.15 p.m., cooking time for the dish is 1 hour but it is not necessary to be cooked in a continuous way. The system can decide to power-on the microwave oven for a period of 45 minutes in the morning when energy is cheaper and then complete cooking time at 7.00 p.m. for a period of 15 minutes right before John comes back home.

3.2 Operator energy services

Rose is not happy at all with the electricity invoices she recently received that are quite high. She would like to understand why his home is so energy consuming and how she can reduce energy bills. She has already subscribed to the services of a communication operator and she can access to the operator web portal. The portal allows her to subscribe to an energy consumption monitoring service.

Once the subscription process is completed, she receives a pack with one or more EMD. If this is her first subscription to operator's home services, she also receives a AIM gateway to be connected to the broadband access network.

After self and simple installation procedure, she can access her personalized service pages on the portal and get a detailed description of energy usage (power and money) of each device with a statistics per device type and period of time. She can also get specific suggestions on how to improve energy usage.

Rose soon realizes the energy services can really provide her some benefits and decides to upgrade her subscription to the premium service. Now she can download a widget from the operator to configure the service and send commands or receive status from the devices. She can also use a dedicated application or an Instant Messaging client on her mobile to access the service. Rose decides to configure the service so that the operator can detect when she is not at home and minimize energy consumption and send her notifications or alarms on the mobile phone.

3.3 Utility services

The utility AimEnergy wants to introduce an incentive based service that allows customers to directly participate in savings and benefits the utility can generate through a flexible cost model. To implement such a business model a communication path between the utility and the customers, which provides the required information (tariff/pricing information), has to be established and the customer needs to have a device which is at least capable of displaying the actual tariff-information. This functionality is provided by the AIM gateway.

For introducing the flexible tariffs, the AimEnergy provides a basic service that sends every day to the user a pricing profile for the next day that includes the energy prices per day hour. The utility can also provide an advanced service that allows customers to buy a certain amount of kilowatt-hours for consumption. This service requires more information exchange through the AIM gateway between the utility and the customer, since a purchase order must be carried out as well as the crosscheck of whether the purchased amount of energy is already consumed.

Aim Energy is also offering to a selected set of users an innovative service that is based on the remote control of energy load. Stephan, one of these customers, owns a special refrigerator which is capable of producing cold air and stores it in a separate part of the fridge that can be released when required. The required amount of energy for that operation is known to the utility and when alternative energy generation is able to provide the needed energy (e.g. strong winds for the windmills) the fridge is remotely activated. This service is considered very useful by both Stephan and AimEnergy since they save money and use more efficiently clean energy sources. A similar service is adopted by AimEnergy to control distributed energy generation creating a kind of virtual power plant that manages the available distributed energy generation sources from the customers connected to its distribution network.

4. USER INTERACTION WITH THE SYSTEM

The ability of the AIM system to adapt to user specific requirements and preferences is a fundamental feature that may determine the level of satisfaction of users and the overall success of this kind of energy management systems.

It is commonly recognized that there are basically two important issues for user acceptance, namely the perception that the system is under direct control all the time and that it is easy to use and able to adapt automatically to needs without complex configuration processes. In the AIM system the user direct control of the system is guaranteed on one side by the possibility to interact manually with all appliances and devices at any moment and on the other side by the definition of a set of user preferences that allow enforcing some specific behaviour. This is obviously not enough for providing an easy to use system, reason for what the AIM system includes also user profiling in order to self-adapt to user habits and to the normal way it uses home appliances. User profiling can take advantage of the supplementary functions provided by a sensor network that can provide some inputs to the system on user identification, user presence at home (and in a specific room), and the level of some physical parameters like the temperature and the light.

4.1 User profiling

User profiling process includes basically two functionalities: a mechanism for recording some events that can characterize the way in which users interact with the home environment and the available appliances, and a simple learning algorithm that allows extracting from all these data some reasonable settings of the energy management system that is expected to be the most appropriate meet user requirements. The final goal of the user profiling function is replacing some of the required system settings based on a manual interaction with a user interface with an automatic configuration procedure that can be performed on request. To this extent, this function must be able to provide inputs to the energy management system exactly in the same way a user could do through the user interface and it can be considered a plug-in of the system that can be enabled of disabled by the user.

The event recording system allows storing the presence of users at home and in specific rooms and the period of times in which it used specific devices according to: i) the day of the week (typically week-days, week-ends, holydays), ii) the time of the day (the granularity may be quite coarse like e.g. half-hour or hour).

From these data the learning system can extract some characteristics of the user habits in the form of probability distributions. For example, it can derive the probability distribution of the user presence at home during week-days and weekends, the probability distribution of the presence in the living room, the probability distribution of using the HiFi audio system, etc. If relevant, the learning system can also extract joint probability distributions like for example the probability of using two devices at the same time.

Based on this user profile characterization, the system can take same decisions on the energy management settings for basically two main purposes:

- Set in a low power mode devices when the probability of being used is very low and set them in some active mode when the user will likely use them (like e.g. activate screen, remote control, etc.);
- Schedule activities requested by the user in periods of times that fits requirements (like e.g. run the washing machine program before the user is back home, heat the milk 5 minutes before the user comes in the kitchen, etc.) .

In some scenarios, user profiling function can take advantage of some feedbacks provided by the user on some undesirable settings performed by the system automatically. For example, if the user is forced to activate manually a device that was set in power save by the system, this can generate a penalty to the learning algorithm that can be translated into a modification of the parameters that determine algorithm decisions.

4.2 Sensor network

The sensor network provides the basic tools for gathering the information on user behaviour and its interaction with appliances from the home environment. Moreover, the sensor network provides measurements of some physical parameters like temperature and light that can be used by the system to perform some automatic adjustment of the energy management system (like e.g. regulating lighting system according to the level of natural light from windows, control the heating/conditioning system to set temperature in the rooms according to the user profile, etc.). The sensor network can also provide a mechanism for user identification (so that different profiles can be created for the different users living in the same apartment/house).

The sensor network can be implemented using several available technologies. However, wireless sensor networks are today considered the most promising and flexible technologies for creating low cost and easy to deploy sensor networks in scenarios like those considered by AIM project.

The sensor nodes can be equipped with several sensing devices. For the AIM scenarios the most relevant sensing devices include presence detection (that can be simple radar based devices or sophisticated localization systems), user identification (like e.g. RFID readers), temperature and light sensing.

Data collected by the sensor network are delivered to a sink node that is in charge of aggregating it and providing inputs to the user profiling module.

5. CONCLUSIONS

In this paper we have presented the main characteristics of the energy management system that is currently being defined by the AIM project.

The final result provided by the project is a system with enhanced home network architecture, incorporating services for home appliances energy consumption monitoring and management and using a generic technology frame so as to be in order to be applicable on other appliance types, e.g. heaters and solar panels. The gateway architecture is, as it had to be, able of performing energy management of home appliances via the home network. The system is a high technology product to be massively adopted by: residential users, for making optimal use of energy at home; service operators, for the development of a new breed of energy aware services for residential use; and power distribution network operators, for optimising their energy generation planning and administer efficiently cases of energy over-demand.

ACKNOWLEDGEMENTS

The authors wish to thank to AIM Consortium and project participants for the important discussions and contribution and to the European Commission for the financial support of AIM project, contract no. 224621: AIM Architecture for Reduction of Energy Consumption of Home Appliances, ICT-2-6.3-FP7.

REFERENCES

The AIM Consortium. A novel architecture for modelling, virtualising and managing the energy consumption of household appliances-AIM, *http:// www.ict-aim.eu*, 2008.

The ESTIA Consortium, Enhanced Networked Architecture for Personalised Provision of AV Content within the Home Environment, *http://www.ist-estia.org/*, 2006.

The European Commission, Action Plan for Energy Efficiency: Realising the potential. Communication from the Commission, Brussels, 2006.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Integrated energy spatial planning: "spatializing" policy decision support

T. Blaschke[a,b], **M. Biberacher**[a], **S. Gadocha**[a], **I. Schardinger**[a], **T. Prinz**[a]

[a] *Research Studios Austria, Research Studio iSPACE, 5020 Salzburg Austra (thomas.blaschke@researchstudio.at; markus.biberacher@researchstudio.at; sabine.gadocha@researchstudio.at; Ingrid.schardinger@researchstudio.at; thomas.prinz@researchstudio.at)*

[b] *University of Salzburg, Z_GIS Centre for Geoinformatics*

Abstract: One objective of a shared environmental information system for Europe (SEIS) is to organise the vast array of already collected environmental data and information and to integrate it for wide spectra of applications and user access. SEIS should pave the road to move away from paper-based reporting to a system where information is managed as close as possible to its source and made available to users in an open and transparent way. In this paper, we report on attempts to use GIS technology and geoportals as means to make spatial information on energy supply, energy potential and demand tangible for spatial planning beyond the provision of static maps. We demonstrate that GIS-based scalable and flexible information delivery sheds new light on the prevailing metaphor of GIS as a processing engine serving needs of users more on demand rather than through 'maps on stock'. We report briefly about several energy related projects in Europe and demonstrate that data will become available together with tools that allow experts to do their own analyses and to communicate their results in ways which policy makers and the public can readily understand and use as a basis for their own actions. Geoportals in combination with standardised geoprocessing today supports the older vision of an automated presentation of data on maps, and – if user privileges are given - facilities to interactively manipulate these maps.

Keywords: Energy modelling; energy region; GIS; spatial planning; GeoPortal; SEIS

1. INTRODUCTION

A reliable, secure, efficient and environmentally sound energy-supply is essential for a sustainable provision of goods and services [de Fries et al. 2007]. Policy is confronted with the challenge of security of supply which is of broad interest, which is a multi-faceted and multi-scaled issue and which needs long term solutions. Improvements of current energy systems concerning CO2 and security of supply are particularly determined by spatial questions. So far, the energy industry has paid only little attention to geospatial aspects in modelling possible future energy systems and solutions. Blaschke et al. [2008] have pointed out the importance of spatial distribution of renewable energy carriers and possible utilization for the energy system. In addition, spatial planning in most European countries is – with exceptions at the local level - not explicitly dealing with "energy spaces", e.g. reserving space for future energy corridors and for "space-consuming" generation of renewable energies such as biomass production.

The problem faced in this context is the generally low energy density of renewable energy carriers which requests more emphasis on geographical deviations of renewable energy supply and energy demand. Although the utility providers are using GIS systems in a very severe manner at central locations within their business chains they are so far mainly thinking “along lines”. In order to reduce the increasingly problematic dependency on fossil fuels, national and regional policies need to take on responsibility for securing their energy supply. Master plans and decisions must be based on hard facts, many of which can and should be based on geographic footprints and on geospatial techniques. Renewable energy sources are characterised by their marked temporal and spatial variability, in contrast with the distribution of the so-called fossil fuels. Typically, one can find at least one local source of renewable energy at almost every location, but this advantage of the broad spectrum of renewable sources compared to the conventional sources, also complicates the energy system. Conventional energy systems are characterised by a concentrated generation model and major consumption points situated far away from resources and power generation.

Renewable energy has a more heterogeneous spatial distribution and tends to be less ubiquitous. Consequently, regional aspects of energy distribution are reinforced today: regional planning will need to explicitly consider and account for various combinations of renewable energy generation, taking into account a region's characteristics and needs in relation to its energy potentials. Blaschke et al. [2008] claimed that only very few studies explicitly deal with geospatial relationships in modelling energy demand and supply [e.g. Hoogwijk et al. 2004, Dominguez & Amador 2007; de Vries et al. 2007, Ramachandra & Shruthi 2007; Pichugina et al. 2008] and hypothesized that this field of overlap between GIS/GIScience and energy research is not well developed yet. Only a few studies connect the supply or the potential supply with the demand side pattern of population [Blaschke et al. 2003] or with the locations of major industry as demand hot spots.

The variability and complexity of energy supply and demand systems necessitates the use of geospatial tools. In this paper we report on a national study in Austria, competed in November 2008, which analyses policy options for integrated energy planning: we elucidate preliminary results of this modelling and, finally, we point out that the internet as such and geoportals in particular enable mediation and facilitate communication for planners and decision makers.

2. RENEWABLE ENERGY MODELLING WITH GIS

2. 1 The general top-down modelling approach

Biberacher [2007] and Biberacher et al. [2008a,b] presented a top-down modelling approach for renewable energy source (RES) potentials. GIS is especially useful in the RES modelling, particularly because of the special geographical aspects of RES. As a first step, universally valid fundamentals are used to calculate the theoretical potentials. The estimation of spatial differentiated theoretical potentials is based on data on topography, climate, land use and many others. The estimated theoretical potentials are then reduced to a technical potential by taking into account technical limitations of state-of-the-art technology as well as factors concerning distribution or topography, e.g. steep slopes. For instance, certain land use classes or protected areas will typically be excluded. By using rather soft factors which may be modified over time and may vary regionally the potential can be further reduced to a realisable one. Under expert-defined assumptions the development and deployment of the individual energy sources are integrated within this step (Figure 1).

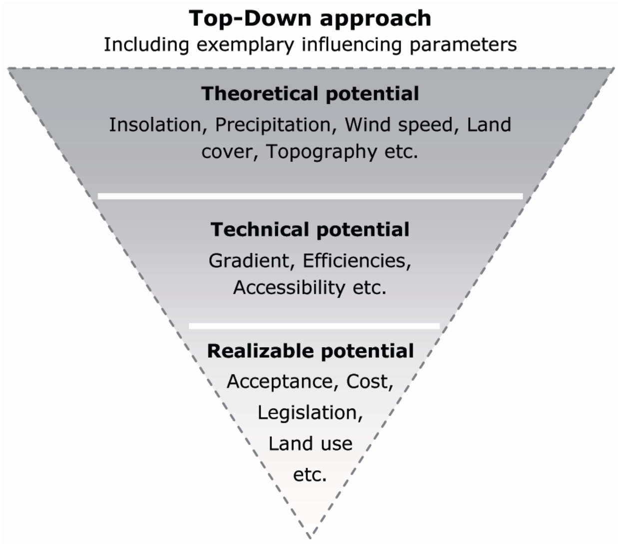


Figure 1. The Top-Down approach in spatially explicit energy modelling.

In addition to energy resources energy demand is assigned to specific locations and energy consumption is modelled at the same geographical resolution as the energy potentials. For the estimation of heat and electricity demand, characteristic values of demand structures are either used directly or are broken down into the appropriate spatial units through disaggregation. Some other statistical data for households in the area of interest are being used for the estimation of energy demand. By combing these data the spatial distribution of the energy demand can be identified and mapped. Figure 2 exemplifies three levels of energy demand presentation for a part of the city of Salzburg, Austria, reaching from address data to census information. The most flexible way to aggregate and disaggregate between the various levels is trough raster representations.

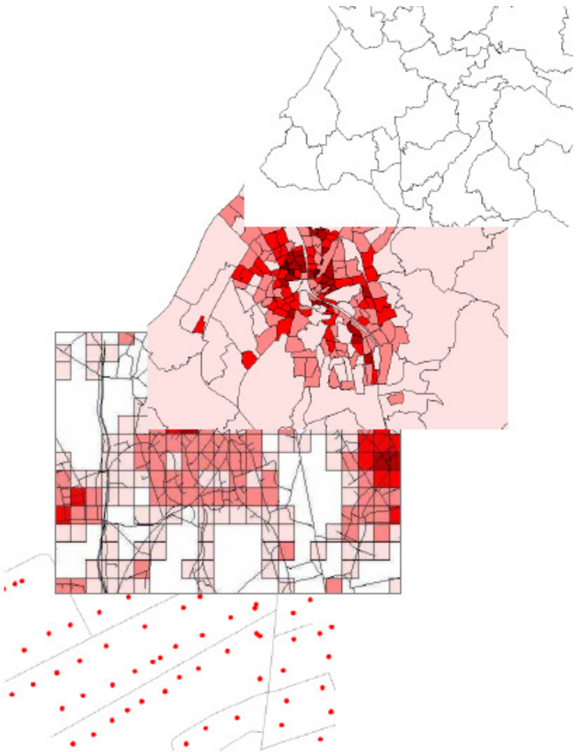


Figure 2. Various levels of representation of energy demand from household data to census information.

Biberacher [2008] optimized the model and further elaborated the framework incorporating location related temporal characteristics in energy supply and demand. These characteristics in mind an imaginable energy system setup can be explored using this framework. In this study the possible coverage of the global energy demand, by solar- and wind power in junction with a backup technology was treated.

2. 2 Earlier studies

Within an Austrian nationally funded research project the concept of Virtual power plants has been implemented based on GIS-technology (Mittlböck et al., 2007). The objective was and still is to create energy self-sustaining regions based on the optimal combination of different renewable energy potentials into virtual power plants and their correlation with the relative energy demand structure. The Research Studio iSPACE has developed an operational method for spatial analyses based on *Regional Autarky Models for a Sustainable Energy Supply* (RAMSES). Biberacher et al. [2008a,b] describe a linear optimisation model and formulate the underlying optimisation problem in the algebraic modelling language GAMS – General Algebraic Modelling Systems and the cplex linear optimisation solver developed by ILOG. The output of an optimisation process yields the arising energy flows and their amount.

Several detailed studies on single energy carriers such as wind, biomass, photovoltaic, solar thermal, geothermic, water etc. have been carried out over the last years in the research studio iSPACE. The level of detail needed in modelling supply, demand and potentials shall hereafter be illustrated for biomass (for more details see Schardinger et al., 2008). In a first instance the biomass to energy service chain may be represented by biomass supply, intermediate storage, final storage, conversion of biomass into usable solid or liquid biofuel, delivery to the energy plant and, towards the end of the chain, energy production. Each link of this chain is more or less directly dependent on the geographical location of each process and will be represented by GIS data sets. They can be output to maps or input to further analyses steps. It is well known that if each step of the whole bio energy chain is not optimised, the final costs of the energy produced (both in terms of heat and electricity) will be high and will be outcompeted by energy from traditional fossil fuels [see, i.e. Perpina et al. 2008].

The cost of energy from biomass mainly results from biomass production, storage, transport from the production or collection area to the plant and plant costs. Moreover, in some cases mechanical, thermo chemical or bio-chemical processes are required to transform the original biomass into the desired more practical fuel after the collection. As a first step of the work the present biomass availability per territory has been determined on the basis of a GIS analysis performed on layers and associated databases based on land use and divided into agricultural resources, forestry potential, green garbage and deposits from food production. At the end of this process biomass potentials are localised, together with information on their typology and the accessibility of the production area.

The modelling of the energy demand follows a bottom-up approach as illustrated in Figure 1: GIS-derived information on settlement structures, population statistics and specific data on consumption patterns of households and economic sectors renders possible to assess the spatial allocation of the energy demand. The comparison of the estimated energy potentials with the existing energy consumption structure results in an 'energy balance map', representing the energy excess or shortage for every sub-unit (Figure 3).

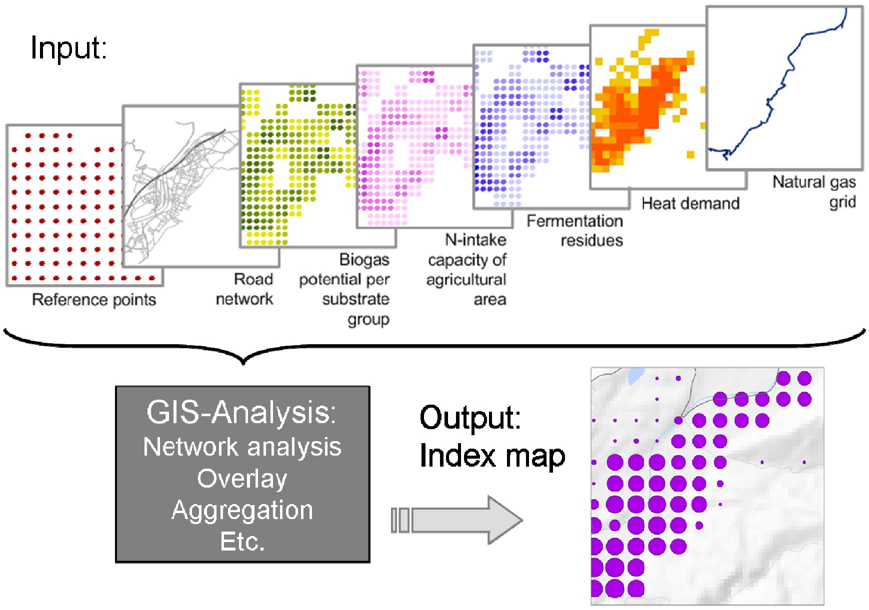


Figure 3. Model for province wide location evaluation for biogas plants

3. ENERGY REGIONS

As well known in scientific literature and in practical spatial planning has to be both top-down and bottom-up, both urban and rural. Basic concepts cannot be repeated herein. It should only be emphasized that spatial planning tools shall flexibly cater for different types of regions (administrative vs. functional e.g. concerning energy demand and supply). It is well know in various business fields that successful regions attract a mix of entrepreneurs, capital, facilities and workers, needed to drive the economy and provide the mix of amenities critical to living the good life. It seems to be less clear if such a combination can be naturally provided concerning energy demand and supply and if those can be usually provided by national and provincial jurisdictions. In spatial planning, the concept of a region is typically seen as a precondition for assessing spatial and regional reference of economic activities when trying to achieve sustainable effects. In the scientific and the political debates a broad variety of regional concepts are in use. The concepts used correspond with heterogeneous ways of a regionalization and its specific and different demands, methods and instruments used for working on regional challenges. The term ‘energy region’ is used pragmatically herein. It is not a question of small-scaled or relatively closed economy, but of a balanced combination of regional economic relationships and of over-regional one's.

An energy region is a hybrid concept, catering for different intentions, partially moving in opposite directions. It includes endogenous regional potentials and it could also be debated in the light of "new regionalism". In spatial planning, these discourses are the departing points of conceptions of "sustainable regional development". Views and findings corresponding with "sustainable regional development" support a conceptual basis for regional approaches to sustainable economy, especially when its spatial relations and effects are regarded. Ultimately, direct connection of the socio-economic actors to regulating subjects on the regional level (e.g. a regional government) will assure regional and sustainable effects of their economical activities. In general, regions are often the most local level for delivering cost-effective water, sewer, transit, and other public services. They have also become the most appropriate level for addressing various challenges. “Energy regions” seem to be a new dimension in defining regions functionally. Generally, they should be large enough to encompass locally/regionally important economic marketplaces and environmental watersheds/airsheds, as well as relatively (!) homogeneous living conditions. Yet energy regions should be small enough to engage community leaders and citizens in pursuing practical strategies to provide affordable infrastructure and services, protect threatened environments, and foster robust economic development aiming to care for local resources without necessarily being fully autarchic.

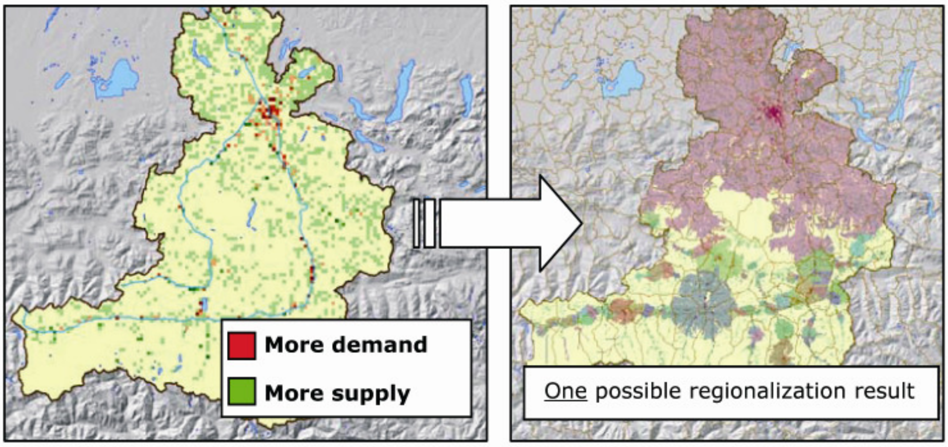


Figure 4. Delineation of energy regions (adapted from: Mittlboeck et al. 2007)

4. RESULTS

For the renewable energy sources hydropower, wind, biomass, solar and geothermal energy, the described general model has successfully been instantiated and applied to regional systems in Austria (Salzburg province and German parts of the Salzach river catchment) and to parts of North-Rhine Westphalia. Through an assessment of spatial interrelations RAMSES supports the decision-making for the adaptation of energy systems. A third study for the whole territory of Austria on behalf of the ÖROK (Austrian Conference on Spatial Planning, an organisation coordinating spatial planning between the federal, province and municipality levels) is a first consensus based process strongly building on GIS-functionality for an integrated spatial planning by outlining renewable energy potentials with their spatial distribution and taking into account spatial planning aspects. (Figure 5).

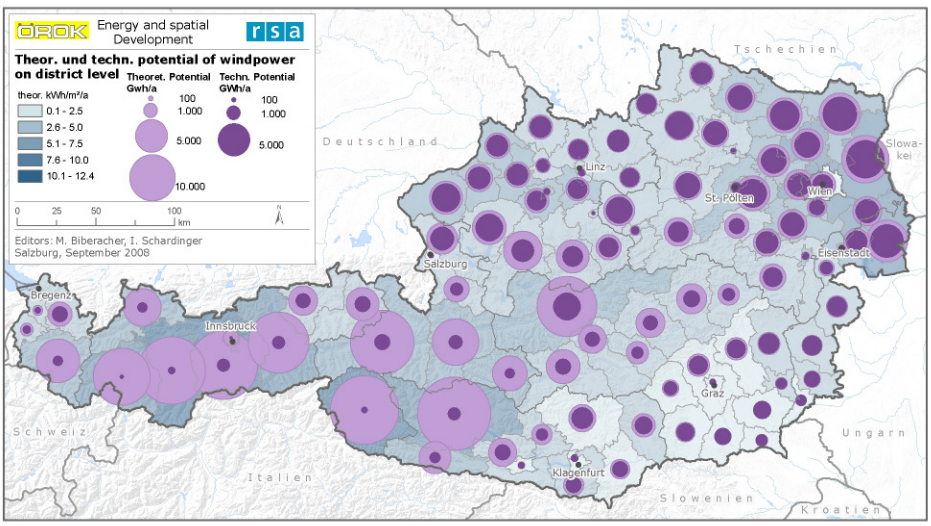


Figure 5. Map of theoretical and technical potential of wind power as one preliminary result from the ÖROK study

The GIS based models serve as the core functionality. The main innovation in the context of integrated systems such as SEIS and geoportals are automated mapping procedures coupled with scenario techniques. The GIS-derived results are typically stored in raster formats and are to be combined further based on the rule sets of the respective user. This way, we incorporate some generic combination functionality (e.g.: potential per cell: {energy carrier

1: 70% realized} + {energy carrier 2: 60% realized} + {energy carrier 3: status quo extrapolated minus estimated demand for time z1}) into a geographically enabled system.

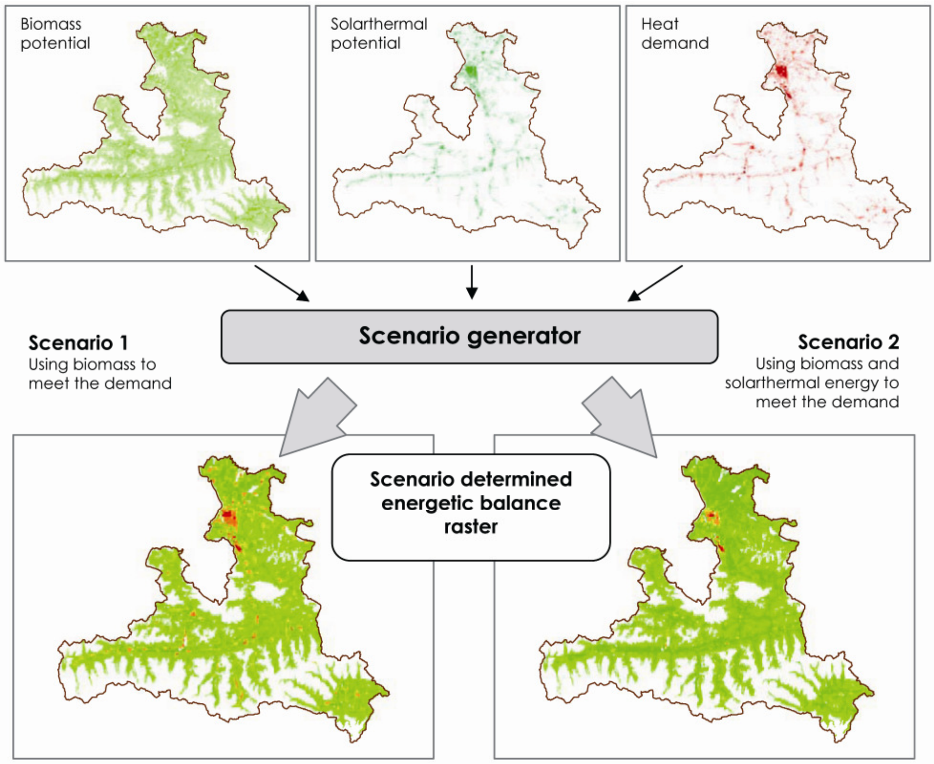


Figure 6. Outlook on further use of results from the ÖROK study for the province of Salzburg. The GIS model shown serves as a scenario generator and communication tool.

We have been reluctant to call this application a 'spatial decision support system' since it has not been tested operationally. If the user has a sound hypothesis and provides logic combinations of expected future energy use the results will be scenarios rather than maps in a more classical sense (see next section). Geographic information based combination allow non-experts to receive proper data presentations, and the automation of map construction saves time. At present, it is not interactive exploratory data analysis with hundreds try and error map generation steps, mainly due to the amount of data and processing steps behind. But the provided interactive, dynamic data displays allow for some degree of interactivity that could enhance the expressiveness of maps and thus promote data exploration.

5. DISCUSSION

As stated at the EC SEIS homepage (http://ec.europa.eu/environment/seis/how.htm), SEIS seeks to take advantage of the opportunities offered by the latest developments in information and communication (ICT) and geographic information (GIS) technologies. The European INSPIRE Directive (2007/2/EC) establishes an infrastructure for spatial information in Europe entered into force in May 2007. It is the first supra-national legal framework building on Geographic Information. It contains provisions aiming to improve the accessibility and interoperability of spatial data. Both, the implementation of INSPIRE and SEIS implementation aim to overcome existing inefficiencies relating to the usability and use of spatial data stored by public authorities.

Since one of the main ideas behind SEIS and legislations like INSPIRE is that data should be collected once and then re-used in different contexts some research questions of Geographic Information Science, or GIScience in short, are tackled. GIScience more and more moves hundreds of applications onto a bigger agenda. The term *GIScience* was coined or at least introduced in the scientific literature by Goodchild [1992], who described it to deal with the basics of GIS-technology, concentrating on those issues that are an impediment to a successful implementation. Duckham et al. [2003] stated in a book on GIScience that GIScience addresses the fundamental research principles on which Geographical Information Systems are based (e.g. research on GIS) or that it refers simply to the use of GIS in scientific applications (e.g. research with GIS). Following the GIScience research agendas discussed by Mark [2003], the authors conclude that concepts of energy regions can be directly or indirectly seen as an important topic for future GIScience research. As Blaschke et al. [2008] claimed energy is at the moment not playing a particular role towards a broader conceptual grounding of successful applications. While early GIS are said to be successful in routine and blunt-edged senses [Longley et al. 2005] they may have shied away from many of the bigger questions concerning how the world works. In tendency, GIS has always been an applications-led technology, and many applications have had quite modest goals in terms of the science of problem-solving. Longley et al. [2005] point out that the test of good science and technology lies in its usefulness for exploring the world around us. Future integrated systems such as SEIS and portals such as GEOSS need to consider that no scientific and technological ingenuity can salvage a representation that is too inaccurate, expensive, cumbersome, or opaque to reveal anything new about the world. Yet GIS applications need also to be grounded in sound concepts and theory if they are to resolve any but the most trivial questions. Kraak [2006] argues that although the traditional paper map allowed geographers to use it to synthesize, analyse and explore, it is obvious that the rise of Geographical Information Systems have stimulated these functions. Maps that used to be produced in an elaborated way can now be created in many alternative views by a single click of the mouse. Additionally, many more maps are produced and used, a trend multiplied by the development of the Internet, and, increasingly, mobile applications.

GIScience is said to be a legitimate subfield of information science and particularly attractive to information scientists because of the well-defined nature of geographic information and the comparatively advanced state of knowledge about this information type [Goodchild 2004]. What are energy regions, then? For renewable energy we have pointed out coarsely how powerful GIS technology is in the modelling part and how important the GIScience research agenda will be in the future in asking the right questions: how can geographic knowledge and skills be acquired and how can concepts of (energy) regions be made more readily understood and usable by humans? It also leads to the question whether or not spatial arrangements of land use types add specific qualities beyond statistical measures of their existence and quantity. For instance, can a landscape be sustainable, as long as 20% of the land use is extensive, 10% is protection area, etc., no matter where the respective patches are, which typical size and shape they have, how connected patches are and how often incompatible land use types are adjacent? [Blaschke 2006]. These questions have to be addressed to energy regions and have to be examined empirically. Translating them into the GIScience world we may follow Gahagen's [2005] dichotomy when developing visually centred methods and techniques and/or tools to present, analyse, synthesize and explore geospatial data but being at the same time interested it their effect on problem-solving (efficiency, effectiveness). Gahagen described the GIScience process, and projected possible maps and graphics, as well as computational methods on each of the process steps. ''To better support the entire science process, we must provide mechanisms that can visualize the connections between the various stages of analysis, and show how concepts relate to data, how models relate to concepts, and so forth'' Gahegan, 2005, p. 85. That is so far not the core vision of worldwide integrating information systems such as SEIS but these are maybe the research questions of tomorrow.

REFERENCES

Biberacher, M. 2007. Modelling and optimisation of future energy systems. Using spatial and temporal methods. VDM Verlag Dr. Müller.

Biberacher, M. 2008. GIS-based modeling approach for energy systems. International Journal of Energy Sector Management, 2(3) 368-384.

Biberacher, M. Gadocha S. and D. Zocher 2008a, GIS based Model to optimize possible self sustaining regions in the context of renewable energy supply. In: M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.), Proc. iEMSs 4th Biennal Meeting. 1306-1313

Biberacher, M., Gadocha, S., Gluhak, S., Dorfinger, N., and D. Zocher 2008b. Implementation of a GEOdatabase to administrate global energy resources. In: M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.), Proc. iEMSs 4th Biennal Meeting. 1223-1230.

Blaschke, T. 2006. The role of the spatial dimension within the framework of sustainable landscapes and natural capital. Landscape and Urban Planning, 75(3-4), 198-226.

Blaschke, T., Mittlböck, M., and G Griesebner, 2003. Development of settlement fragmentation indices for energy infrastructure cost assessment in Austria. Int. Archives of Photogrammetry, Remote Sensing & Spatial Inform. Sciences, XXXIV-7/W9, 24-28.

Blaschke, T., Biberacher, M, Gadocha, S., Zocher, D., Mittlböck, M., Haslauer, E., Schardinger, I. and J. Strobl 2008. Virtual Power Plants: spatial energy models in times of climate change. In: Ehlers, M. et al. (Eds.), Digital Earth Summit on Geoinformatics 2008, Wichmann-Verlag, Heidelberg, 61-66.

de Vries, B.J., van Vuuren, D.P., and M.M. Hoogwijk 2007. Renewable energy sources: Their global potential for the first-half of the 21st century at a global level: An integrated approach. Energy Policy, 35, 2590–2610.

Dominguez, J. and J. Amador 2007. Geographical information systems applied in the field of renewable energy resources, Computers and Industrial Engineering, 52(3), 322–326.

Duckham, M., Goodchild, M.F. and M. Worboys (eds.) 2003. Foundations of Geographic Information Science. Taylor & Francis, London.

Gahegan, M. 2005. Beyond tools: visual support for the entire process of GIScience. In: Dykes, J., MacEachren, A. M. and M.-J. Kraak (eds.), Exploring Geovisualization, Elsevier, Amsterdam, 83–99.

Goodchild, M.F. 2004. GIScience, Geography, Form, and Process. Annals of the Association of American Geographers, 94(4), 709–714.

Hoogwijk, M., de Vries, H.J.M., and W.C. Turkenburg 2004. Assessment of the global and regional geographical, technical and economic potential of onshore wind energy. Energy Economics, 26, 889–919.

Kraak, M.-J. 2006. Why Maps Matter in GIScience. The Cartographic journal, 43(1), 82–89.

Longley, P., Goodchild, M.F., Maguire, D. and D. Rhind 2005. Geographic Information Systems and Science , 2nd Edition, Wiley & Sons, Chichester.

Mark, D.M. 2003. Geographic Information Science: defining the field. In: Duckham, M., Goodchild, M.F. and M.Worboys (eds.), Foundations of Geographic Information Science. pp. 3–18, Taylor & Francis, London.

Mittlböck, M. et. al. 2007. Virtuelle Kraftwerke für Autarke Regionen. (BMVIT, Ed., Berichte aus Energie- und Umweltforschung 58/2006), Wien.

Perpina, C., D. Alfonso, A. Pérez-Navarro, E. Peñalvo, C. Vargas, and R. Cárdenas 2008. Methodology based on Geographic Information Systems for biomass logistics and transport optimisation. Renewable Energy, 34(3), 555-565.

Pichugina Y.L., R.M. Banta, N.D. Kelley, W.A. Brewer, S.P. Sandberg, J.L. Machol and B.J. Jonkman 2008. Remote sensing of the nocturnal boundary layer for Wind Energy applications. 14th Int. Symp Adv. Boundary Layer Rem. Sens., 1-10.

Ramachandra, T.V. and B.V. Shruthi 2007. Spatial mapping of renewable energy potential. Renewable and Sustainable Energy Reviews, 11(7), 1460–1480.

Schardinger, I., Biberacher, M. and S. Gadocha 2008. GIS based Location Decision Support Model for Biogas Plants considering local Supply and Demand. In: Schnitzer, H. and S. Ulgiati (eds.), Advances in Energy Studies. Towards a holistic approach based on science and humanity, Verlag der Technischen Universität Graz, Graz, 456 – 464..

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, J. Pelikán, O. Mírovský, W. Pillmann, I. Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Power Optimized More Electrical Aircraft

Jean-Luc Derouineau

Honeywell Aerospace, Director of European Research and Technology
(Jean-Luc.Derouineau @ Honeywell.com)

Abstract: The last few years Honeywell Aerospace has been developing aircraft power optimization technology in sense of More Electric Architectures studies. Higher electrical load demand required from both defense and commercial applications will influence the configuration of future More Electric Engines (MEE). MEE configurations related studies have shown that MEA systems can be optimized to achieve the highest performance, friendly environmental (more "green") operation (by reducing fuel burn, CO2-emission, air contamination, noise level etc) and lowest weight aircrafts. One of the most promising Power Management System (PMS) developments is a pure "green" fuel cell technology that uses a "cold" combustion process. Some advanced fuel cell technology, such as a proton exchange membrane (PEM) could replace aircraft auxiliary power unit (APU) and emergency power unit likes a Ram Air Turbine (RAT).

Keywords: More Electrical Aircraft; More Electrical Engine; Power Management System.

1. INTRODUCTION

MEA is one of Honeywell's most advanced design concepts for building aircraft of the future. This breakthrough technology, which has major applications for the future of military and commercial aircraft design, eliminates much of the pneumatic and hydraulic fluid and power systems required in today's traditional heavy, maintenance-intensive systems.

By utilizing MEA, aircraft can be designed with lighter and simpler PMS to reduce operating and maintenance costs. On Figure 1, we can appreciate the shift to electric power for some traditional aircraft's functions such as: Environmental Control System (ECS), Ice Protection System (IPS), Flight Control Actuation (FCA) and Main Engine Start (MES).

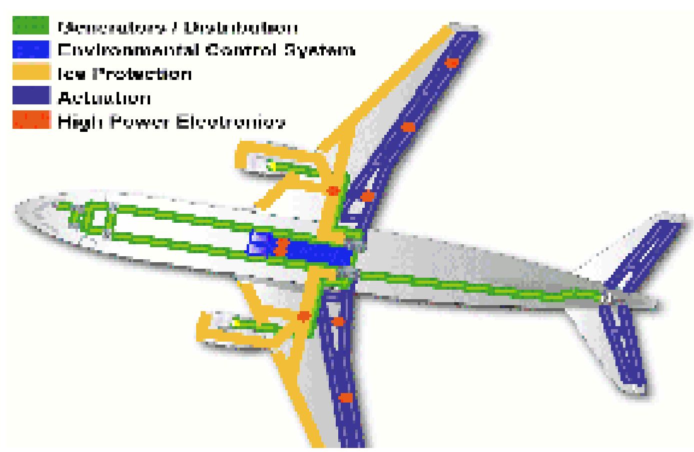


Figure 1. Aircraft systems

The adoption of MEA–concept by military and commercial aircraft will also enable better sensing and monitoring of the systems allowing more reliable detection and prediction of failures. The result will be simplified maintenance, improved dispatch reliability, and more “green” operation.

2. POWER MANAGEMENT BENEFITS

As the next-generation MEA architecture is adopted, electric systems and components will deliver significant benefits industry and environment wide. Honeywell’s MEA solutions reduce fuel consumption, increasing overall aircraft performance and energy usage. Reduced maintenance and ground support are additional benefits to lower cost of ownership and operation. Shifting from traditional pneumatic and hydraulic power sources to electrical power creates significant changes to the electrical system requirements in both sizing and availability. Figure 2 gives typical s electrical load for the systems demand and the corresponding generation requirements.

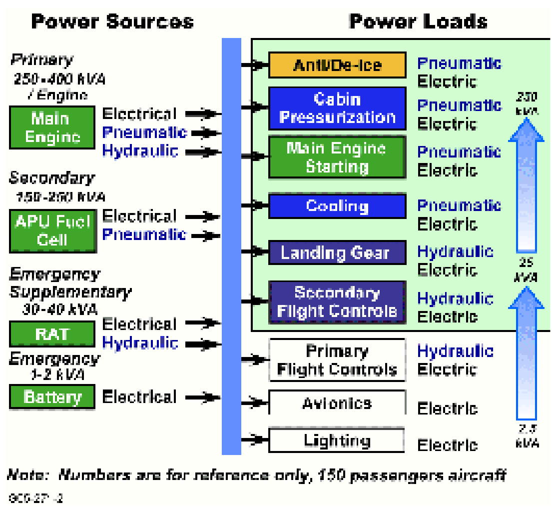


Figure 2. Aircraft systems typical electrical load and generation required

When all aircraft’s systems are designed and optimized for MEA, it is possible to achieve performance improvement during all flight phases compared to traditional PMS architecture. Let’s take for example, just Electrical ECS and IPS benefits:

- Engine power off-takes reduces by a few hundreds %
- Specific Fuel Consumption reduces up to 2%
- Engine thrust increases up to 15%
- Engine operates cooler up to 40 deg C
- Engine and APU contamination reduces by 10-15%

3. MORE ELECTRICAL ENGINE (MEE)

In order to achieve all these benefits, engines and their accessories have to be designed with different requirements in mind: a lot of the components and associated ducting and piping that traditionally generate the complexity of the nacelle installation can be significantly simplified and streamlined when going MEE.

Some sources of heat and flammable liquids are eliminated that simplifies further installation and certification. As more and more of the mechanical accessories are converted to electrical, it is possible to envision the simplification or the suppression of the engine gear box that opens the route to “gear less” and eventually “oil less” engines.
Figure 3 provides a comparison of a traditional engine installation and an MEE installation.

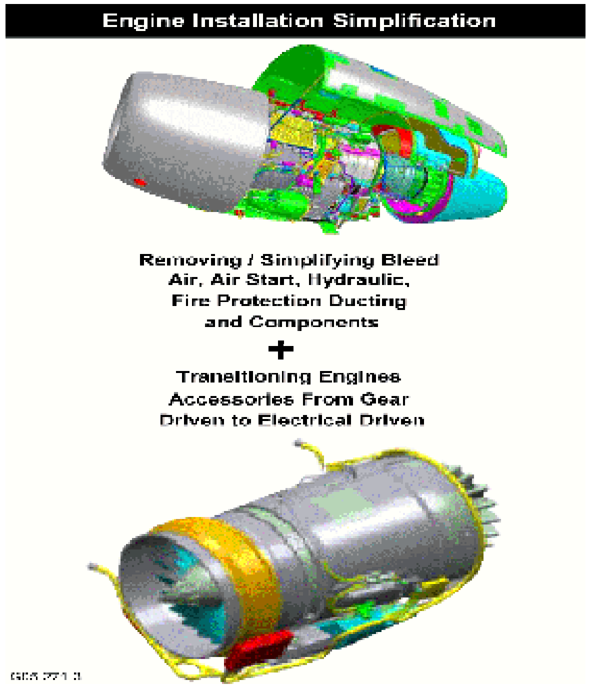


Figure 3. Transition in installation from conventional engine to MEE

For a new dedicated MEE cycle the operating line can be placed in the most advantageous location. Extracting power from the lower pressure shafts provides the better climb in icing conditions and alleviates enrooted limitations when one engine is inoperative.

4. SUMMARY

The future of MEA technology maturing continues to open opportunities for aircraft PMS improvements. MEA technologies are continually evolving and have huge potential for improvement in weight, volume and cost at the component level while other traditional systems are on the asymptote of their maturation curve. Even more important is the fact that MEA is a “must have” enabler for emerging technologies like: hydrogen power and fuel cells, which are one the few game changing opportunities in Aeronautics horizon.

REFERENCES

Jean Luc Derouineau, More Electrical Architectures and the Impact on Propulsion Systems, 17th International Symposium on Air Breathing Engines, ISABE, 2005, Munich, Germany

Philip Butterworth-Hayes, All-electrical Aircraft Research Speed up, Aerospace America Journal, January 2009

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Modelling of the Taylor-Green Vortex by the Implicit Large Eddy Simulation

V. Fuka [a], **J. Brechler**[a]

[a]*Dept. of Meteorology and Env. Protection, Charles University, V Holešovičkách 2, Prague 8, 180 00, Czech Republic (vladimir.fuka@gmail.com, josef.brechler@mff.cuni.cz)*

Abstract: Implicit Large Eddy Simulation is an emerging tool for turbulence modelling. Instead of explicitely computing subgrid stress models, numerical dissipation of nonlinear schemes can be used as a main control mechanism of turbulent energy transfer. In this contribution we test performance of a high-resolution incompressible projection method in modelling of the Taylor-Green vortex flow, which stands as a prototype of a simple free flow with a transition to turbulence and a turbulence decay. Several criteria are used to asses the model, including the kinetic energy dissipation rate, kinetic energy spectra and probability density functions of velocity gradients.

Keywords: large eddy simulation; turbulence decay; turbulence simulation

1 INTRODUCTION

At the present time a new field of turbulence modelling called the Implicit Large Eddy Simulation (ILES) (for a comprehensive review see Grinstein et al. [2007b]) enables turbulent flow computations with methods most commonly used for the compressible fluid dynamics with shocks. These so called shock-capturing schemes can describe flows with very strong gradients using convection terms with nonlinear numerical diffusion. This numerical dissipation is stronger in areas with larger gradients and smaller in smooth parts. Nonlinearity can be achieved using flux or slope limiters [Harten, 1983; van Leer, 1979]. Especially the second way proved to be efficient way of simulating compressible turbulent flows. Because of monotonicity preserving of these schemes, this kind of LES is also known as MILES - monotonically integrated large eddy simulation). To the most used methods belongs flux-corrected transport (FCT) [Boris and Book, 1997] or MPDATA [Smolarkiewicz and Margolin, 1998]. MPDATA is used mainly for geophysical applications, which is also aim of our work.

Projection methods, also known as fractional step methods [Brown et al., 2001], can be divided into two groups – exact projection methods and approximate projection methods. Approximate projection methods mostly employ cell centered grids and enable easier usage of high resolution methods [Almgren et al., 1996, 1998] at the cost of more difficult treatment of velocity – pressure coupling. Exact projection methods [Kim and Moin, 1985] provide very velocity – pressure coupling on staggered grids, but the usage of high resolution methods is complicated. First example of an exact projection method on a staggered grid with a high resolution advective scheme was a method by Tau [1994]. This scheme used a projection method of Bell et al. [1989] with Godunov method for advective fluxes modified for a staggered grid in 2D. We choose this method as a base for our 3D model for incompressible implicit large eddy simulation.

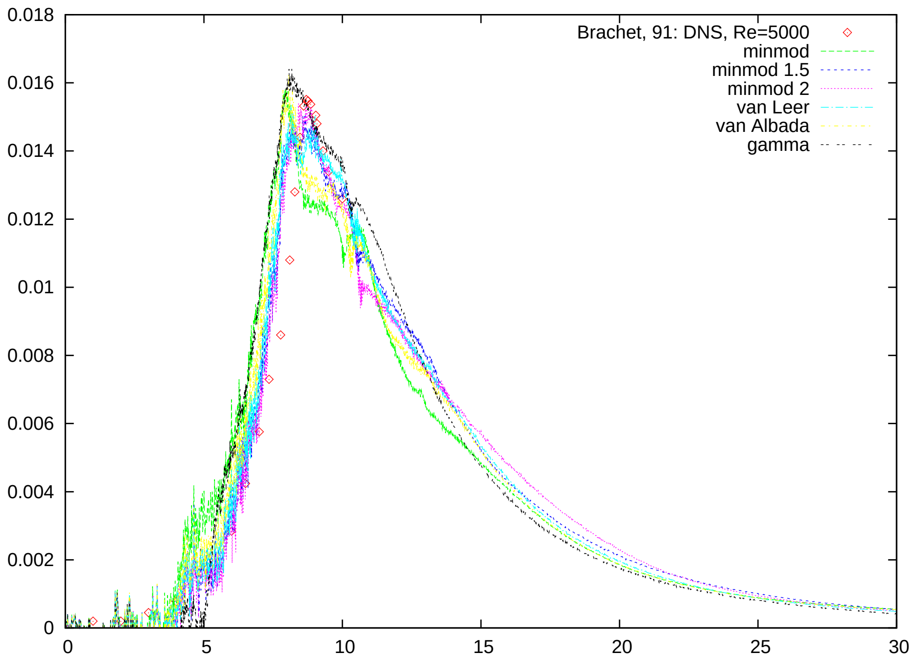


Figure 1: Plot of kinetic energy dissipation at resolution 256^3.

2 Results

As a first test case we computed Taylor-Green vortex flow, which serves as a simple example of a free flow with transition to turbulence and the subsequent turbulent decay. This flow was used for this purpose by several other authors, for example [Garnier et al., 1999; Grinstein et al., 2007a]. The comparison is usually done with the direct numerical simulation (DNS) data of Brachet et al. [1983] and Brachet [1991].

In figure 1 is plotted the time history of kinetic energy dissipation, scaled for a direct comparison with Brachet [1991]. From the plot it is clear that the basic features of the time dependence are captured by the calculations. The differences between individual limiters are visible. The most distinct one is the minmod limiter, which turns out to be overly diffusive. From the other limiters, extended minmod seems to produce result closest to the DNS.

We also computed the probability density functions (PDFs) of the velocity gradients and pressure and the 3D kinetic energy spectra. In figure 3 are plotted examples of the PDFs. For the velocity the tails of the distribution show almost exponential nature, as expected. The 3D energy spectrum (in figure 2) for t up to 8 contain a clear inertial subrange with the Kolmogorov $k^{-5/3}$ power law. The higher wavenumbers are probably affected by the numerical dissipation of the scheme. Similar results were reported by Garnier et al. [1999].

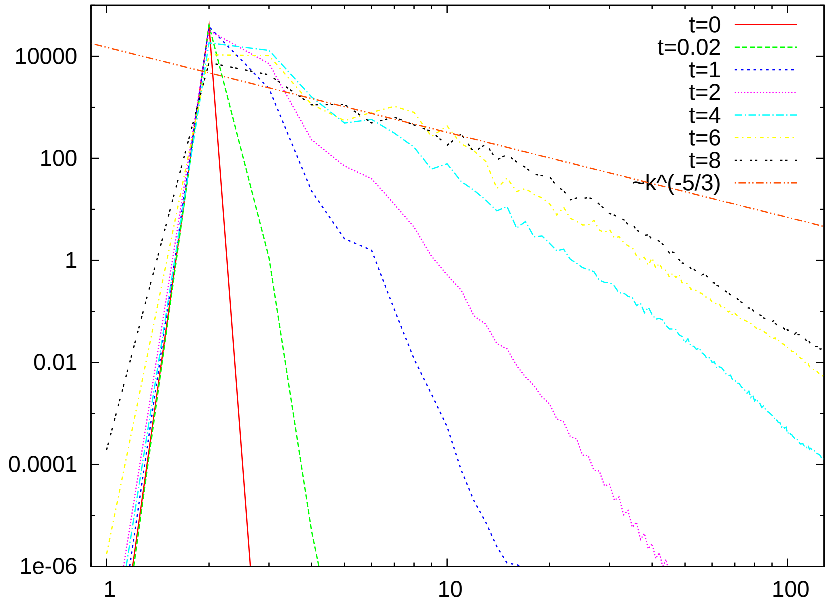


Figure 2: 3D kinetic energy spectra for $t \in [0, 8]$.

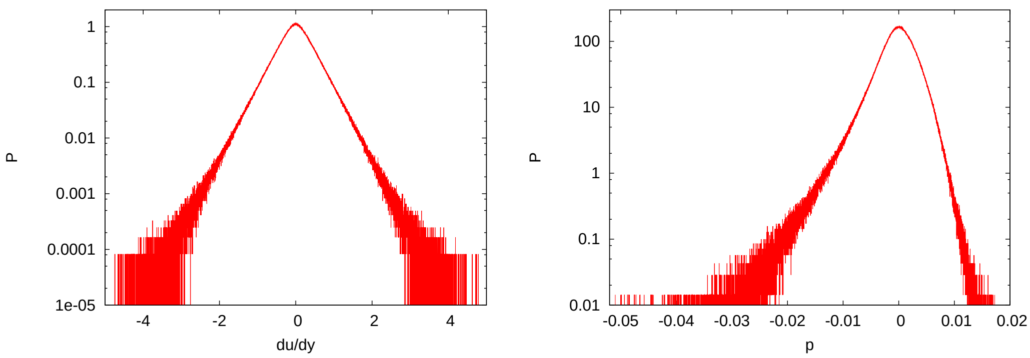


Figure 3: Probability density functions of $\frac{\partial u}{\partial y}$ (left) and p (right).

3 Conclusions

We have developed a 3D model for incompressible flows using a projection method and a high-resolution method for the advective fluxes. We tested capability of this model for turbulent flow as an implicit large eddy simulation model using the Taylor-Green vortex. The model was capable to describe correctly many features of turbulent flow including the kinetic energy dissipation and kinetic energy spectra. The development of the model will continue.

Acknowledgments

This research was supported by the Grant Agency of the Czech Academy of Sciences, grant no. T400300414, by the Grant Agency of the Charles University , grant no. 75908 and by the Czech Ministry of Education, Youth and Sports in the framework of the research plan MSM0021620860.

References

Almgren, A., J. Bell, P. Colella, L. Howell, and M. Welcome. A conservative adaptive projection method for the variable density incompressible Navier-Stokes equations. *J. Comput. Phys.*, 142:1–46, 1998.

Almgren, A., J. Bell, and W. Szymczak. A numerical method for the incompressible navier-stokes equations based on an approximate projection. *SIAM J. Sci. Comput.*, 17:2, 1996.

Bell, J. B., P. Colella, and H. M. Glaz. A second-order projection method for the incompressible Navier-Stokes equations. *J. Comput. Phys.*, 85:257–283, 1989.

Boris, J. P. and D. L. Book. Flux-corrected transport I. SHASTA, a fluid transport algorithm that works. *J. Comput. Phys.*, 135(2):172–186, 1997.

Brachet, M. E. Direct simulation of three-dimensional turbulence in the Taylor-Green vortex. *Fluid Dyn. Res.*, 8:1–8, 1991.

Brachet, M. E., D. I. Meiron, S. A. Orszag, B. G. Nickel, R. H. Morf, and U. Frisch. Small-scale structure of the Taylor-Green vortex. *J. Fluid Mech.*, 130:411–452, 1983.

Brown, D. L., R. Cortez, and M. L. Minion. Accurate projection methods for the incompressible Navier-Stokes equations. *J. Comput. Phys.*, 168:464–499, 2001.

Garnier, E., M. Mossi, P. Sagaut, P. Comte, and M. Deville. On the use of shock-capturing schemes for large-eddy simulation. *J. Comput. Phys.*, 153(2):273–311, 1999.

Grinstein, F., L. Margolin, and W. Rider. chapter Numerics for Iles, Limiting Algorithms. Cambridge University Press, Cambridge, 2007a.

Grinstein, F., L. Margolin, and W. Rider. *Implicit Large Eddy Simulation. Computing Turbulent Fluid Dynamics*. Cambridge University Press, Cambridge, 2007b.

Harten, A. High resolution schemes for hyperbolic conservation laws. *J. Comput. Phys.*, 49: 357–393, 1983.

Kim, J. and P. Moin. Application of a fractional-step method to incompressible Navier-Stokes equations. *J. Comput. Phys.*, 59:308–323, 1985.

van Leer, B. Towards the ultimate conservative difference scheme V. a second-order sequel to Godunov's method. *J. Comput. Phys.*, 32:101–136, 1979.

Smolarkiewicz, P. K. and L. G. Margolin. MPDATA: a finite-difference solver for geophysical flows. *J. Comput. Phys.*, 140(2):459–480, 1998.

Tau, E. Y. A second-order projection method for the incompressible Navier-Stokes equations in arbitrary domains. *J. Comput. Phys.*, 115(1):147–152, 1994.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

High Resolution Regional Climate Change Modelling – Another Challenge for ICT Development

T. Halenka[*)], Michal Belda, and Jiří Mikšovský

Charles University, Prague, Faculty of Mathematics and Physics, Department of Meteorology and Environment Protection, V Holesovickach 2, 180 00 Prague, Czech Republic (tomas.halenka@mff.cuni.cz)

**) Regular Associate of the Abdus Salam International Centre for Theoretical Physics, Trieste, Italy*

Abstract: There is significant problem for decision making process arrising from the weak link between climate change information based on global climate models and impact studies necessarily based on real local conditions. Global Circulation Models (GCMs) can reproduce reasonably well climate features on large scales (global and continental), but their accuracy decreases when proceeding from continental to regional and local scales because of the lack of resolution. This is especially true for surface fields, such as precipitation, surface air temperature and their extremes, which are critically affected by topography and land use. However, in many applications, particularly related to the assessment of climate-change impacts, the information on surface climate change at regional to local scale is fundamental. To bridge the gap between the climate information provided by GCMs and that needed in impact studies, several approaches have been developed. The most popular approaches are (i) statistical downscaling, i.e., identification of statistical relationships between large-scale fields and local surface climate elements, and (ii) dynamical downscaling, i.e., nesting of a fine scale limited area model (or Regional Climate Model, RCM) within the GCM. The latter approach is more correct from a physical point of view, but is much more demanding on computer resources. Similarly as GCMs were and still have been the challenge for the development of ICT with main emphasis to the speed, RCMs become to be problem with rather the storage capacity requirements. A resolution sufficient to capture the effects of topographical and associated land-use features is necessary, that is why 10 km resolution has been introduced in EC FP6 project CECILIA. The main objectives of the project CECILIA dealing with climate change impacts and vulnerability assessment in targeted areas of Central and Eastern Europe are described as well as the first results achieved at Charles University.

Keywords: Climate change; Regional climate modelling; Climate change impacts.

1. INTRODUCTION

Although the broad response of global climate to increased greenhouse gas concentrations is well established, many unknowns remain in the regional details of projections of future climate change. The floods and droughts which occurred in recent summers in the region highlight the importance of the hydrologic cycle and water management in Elbe and Danube river catchments in response to the occurrence of precipitation extremes. In summer of 2002 the Czech Republic experienced some of its worst floods in history with the Vltava river inundating Prague causing severe and widespread damage. The 2003 heat wave, one of the severest heat waves on record in central and western Europe causing both human losses [Kysely, 2004] and extensive damage to human activities and natural ecosystems, demonstrated the importance of the health impacts of extreme conditions that could also lead to considerable changes in air quality, both regionally and in major urban centres. The possibility of changes in the frequency and intensity of these extreme events is one of the most dreaded manifestations of anthropogenic climate change. A number of studies have linked the occurrence of these extreme events to anthropogenic forcings [Beniston and Stephenson, 2004, Schär et al., 2004; Pal et al. 2004; Meehl et al., 2004; Meehl and Tebaldi, 2004]. Impacts on agriculture and forestry affecting the economy of countries in the region are extensively studied as well [Menzel et al., 2006, Menzel et al., 2003, Hafner, 2003, Gobron et al., 2005].

While coupled atmosphere-ocean general circulation models (AOGCMs) provide basic information on the development of the climate change, their horizontal resolution is still too coarse to describe in detail processes affecting extreme events at the regional scale [Gates et al., 1996]. In order to regionally enhance the AOGCM information a number of regionalization techniques have been developed [Giorgi et al., 2001]. One of them is the use of limited area regional climate models (RCMs) nested in driving fields from reanalysis or GCMs similarly as previously done in numerical weather prediction for decades [Giorgi and Mearns, 1999]. The main advantage of this method is that it can reach with the same or even less resources horizontal grid intervals of a few tens of km and thus RCMs can provide improved simulation of extreme events [e.g. Huntingford et al., 2003; Frei et al., 2003]. During the last decade RCMs have been increasingly used to examine climate variations at scales that are not resolved by global models. To the extent that they produce realistic climate simulations, such models can be powerful tools in the study of regional climate impacts.

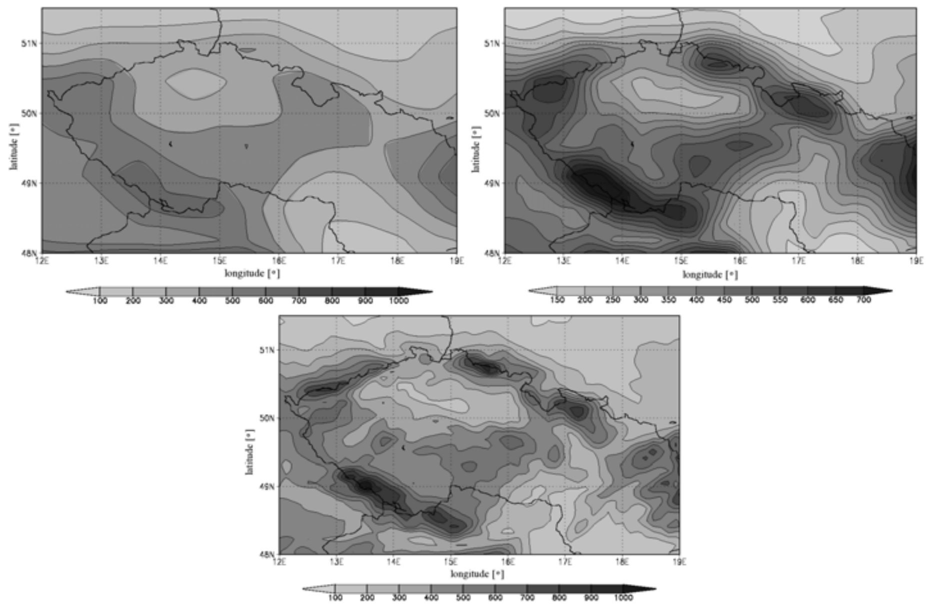


Figure 1. Detail of topographical features seen in ENSEMBLES' 50 km resolution (upper left) and 25 km (upper right) and 10 km for CECILIA proposal (bottom panel).

Thus, the aim of the EC FP6 project CECILIA (**C**entral and **E**astern Europe **C**limate Change **I**mpact and Vulnerabi**li**ty **A**ssessment) is to assess the impact of climate change at the regional to local scale for central and eastern Europe (CEE). using very high resolution simulations in order to capture the effects of the complex terrain of the region. In this region of central and eastern Europe the need for very high resolution studies is particularly important. This region is characterized by the northern flanks of the Alps, the long arc of the Carpathians, and smaller mountain chains and highlands in the Czech Republic, Slovakia, Romania and Bulgaria that significantly affect the local climate conditions. A resolution sufficient to capture the effects of such topographical and associated land-use features is necessary as illustrated in Fig. 1, where comparison of model topography representation in resolutions of 50 and 25 km used for EC FP6 project ENSEMBLES to the resolution of 10 km introduced in CECILIA project is presented in the detailed view on the Czech Republic. Additionally, the central internal objectives of CECILIA are to improve regional climate scenarios and their localization for climate impacts models, and comparing these results against the results of previous and ongoing projects to assess the added value of dynamical downscaling at very fine scales.

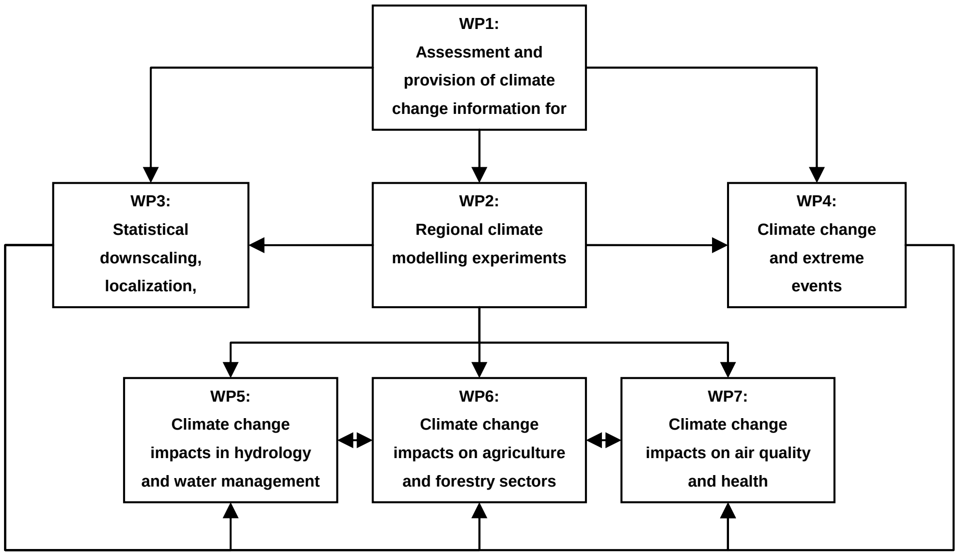


Figure 2. Interactions between the scientific/technical work packages

These goals will be achieved mainly using very high resolution RCMs run locally for individual targeted areas. As climate change scenario projections the production of two time slices are planned, for 2020-2050 and 2070-2100, with 1961-1990 as a reference period. Weather pattern changes as well as changes of extreme events apppearance are addressed within the project as they affect the important sectors of the economies and welfare of individual countries in the region. To address the uncertainties the results of previous or ongoing projects (PRUDENCE, ENSEMBLES) will be evaluated and used here. The selected applications of the CECILIA outputs are supposed toward water resources and management, agriculture, forestry, air quality and health. The objectives will be achieved by means of the following specific tasks (for basic structure with interactions between the individual tasks - work packages, WP, see Fig. 2):

- To collect, assess and make available for first local impact studies the scenarios and climate simulations produced in previous relevant projects, especially PRUDENCE, STARTDEX, MICE and ENSEMBLES, where available. (WP1, see Giorgi and Coppola, 2007)
- To adapt and develop very high resolution RCMs for the region (10 km grid spacing) and perform regional time-slice nested runs driven by ERA40 data and by GCMs for selected GHG change scenarios. (WP2)
- To verify the model results, compare RCM and statistical downscaling results, analyze and develop the methods for verification, particularly at local scales, to provide the scenarios. (WP3)
- To estimate the effect of global climate change on extreme events in the region, including the assessment of the added value of high-resolution for the simulation of the

relevant processes and feedbacks. To evaluate uncertainties in regional projections by comparing results from previous projects (WP4)
• To assess (using high resolution downscaling results) the impacts of climate change on the hydrological cycle and water resources over selected catchments; the effects of climate change on the Black Sea (WP5)
• To study (based on the high resolution downscaling results) the impacts of climate change on agriculture and forestry, carbon cycle and selected species (WP6)
• To study (based on the high resolution downscaling results) the impacts of climate change on health and air quality (photochemistry of air pollution, aerosols) (WP7)

The main goal of the CECILIA project is to integrate results from different previous and ongoing modelling activities and approaches to provide the basis for very high resolution climate change impact and vulnerability assessment in important human activity sectors and natural ecosystems. It is prohibitive to cover within the STREP (Specific Targeted Research Project) all the sectors in their complexity, so that we limited our analysis on some key areas of specific interest to the regions. Nevertheless, the scope of the project is big enough and highly multidisciplinary, which required quite big team. The consortium comprises 16 partner institutions (see Tab. 1) that bring together a wide range of inter-disciplinary expertise and experience in the areas of climate change modelling, statistical analysis, climate change impact assessment as well as anthropogenic and biogenic pollutant emissions. Reflecting the shared and overlapping expertise and interests of the partners, most of the work packages involve contributions from several partners, under the direction of a lead partner. As the targeted area is central and eastern Europe the substantial part of participant institutes comes from this area (10-12, depending on the political or geographical point of view), complemented by selected partners from western part of EU to add necessary access to experience, know how and data for working on the project.

Table 1. List of partners in project CECILIA.

Partner name	Abbreviation	Country
Charles University, Prague	CUNI	Czech Republic
The Abdus Salam ICTP, Trieste	ICTP	Italy
Météo-France, Toulouse	CNRM	France
Danish Meteorological Institute, Copenhagen	DMI	Denmark
Aristotle University of Thessaloniki	AUTH	Greece
Czech Hydrometeorological Institute, Prague	CHMI	Czech Rep.
Institute of Atmospheric Physics, Prague	IAP	Czech Rep.
Swiss Federal Institute of Technology Zurich	ETH	Switzerland
University of Natural Resources and Applied Life Sciences, Vienna	BOKU	Austria
National Meteorological Administration, Bucharest	NMA	Romania
National Institute of Meteorology and Hydrology, Sofia	NIMH	Bulgaria
National Institute of Hydrology and Water Management, Bucharest	NIHWM	Romania
Hungarian Meteorological Service, Budapest	OMSZ	Hungary
Forest Research Institute, Zvolen	FRI	Slovakia
Warsaw University of Technology, Warsaw	WUT	Poland
Eötvös Loránd University, Budapest	ELU	Hungary

2. CECILIA RCMs AND SIMULATION STRATEGY

One of the commonly used RCM in the targeted regions is the model RegCM distributed freely from ICTP. The model was originally developed by Giorgi et al. [1993a,b] and later augmented as described by Giorgi and Mearns [1999], Pal et al. [2000] and Pal et al. [2007]. The dynamical core of the RegCM is equivalent to the hydrostatic version of the NCAR/Pennsylvania State University mesoscale model MM5. Surface processes are represented via the Biosphere-Atmosphere Transfer Scheme [BATS - Dickinson, 1993] and boundary layer physics is formulated following a non-local vertical diffusion scheme [Giorgi et al., 1993a]. Resolvable scale precipitation is represented via the scheme of Pal et al. [2000], which includes a prognostic equation for cloud water and allows for fractional grid box cloudiness, accretion and re-evaporation of falling precipitation. Convective precipitation is represented using a mass flux convective scheme [Giorgi et al., 1993b] while radiative transfer is computed using the radiation package of the NCAR Community Climate Model, version CCM3 [Giorgi et al., 1999]. This scheme describes the effect of different greenhouse gases, cloud water, cloud ice and atmospheric aerosols. Cloud radiation is calculated in terms of cloud fractional cover and cloud water content, and the fraction of cloud ice is diagnosed by the scheme as a function of temperature. For more details on the use of the model see Elguindi at al. [2006].

Another RCM has been used recently, starting from an operational NWP model used in several national meteorological services in the targeted domain. This is the ALADIN-CLIMATE model and first experiences from its development can be found in Huth et al. [2003]. Originally, the limited-area prediction model ALADIN has been developed by the international team headed by Météo-France and modification for RCM purposes started in 2001 in cooperation with CHMI in Prague. ALADIN is a fully three-dimensional baroclinic system of primitive equations using a two-time-level semi-Lagrangian semi-implicit numerical integration scheme and digital filter initialisation. For the description of the model and its parameterizations, refer e.g. to Bubnová et al. [1994] and Váňa [1998]. The physical parameterizations package comprises gravity wave drag parameterization, implicit horizontal diffusion computed in spectral space (fourth order and increasing with height), vertical diffusion and planetary boundary layer parameterization, constant analysed sea surface temperature and amount of sea-ice, an improved version of the ISBA (Interaction Soil Biosphere Atmosphere) scheme, including an explicit parameterization of soil freezing (prognostic variables in ISBA: surface temperature, mean soil temperature, interception water content, superficial soil water content, total liquid soil water content, total frozen soil water content), simple parameterization of snow cover, soil characteristics (texture, depth) that are point-dependent, vegetation characteristics that are point- and month-dependent, simplified radiation scheme called at every time step, mass flux convection scheme including the entrainment profile, specific humidity as a solely prognostic variable: no storage of condensate; evaporation of falling rain; treatment of the ice-phase, and a sophisticated diagnostic cloud (and cloud content) method used for radiative transfer calculations. For running the ALADIN model in a climate mode, a few modifications had to be made, which include mainly changes in lower boundary condition specifications and availability of restart.

Another way to increase resolution is to use GCM with a variable horizontal resolution, e.g. model ARPEGE [Déqué and Piedelievre, 1995; Déqué et al., 1998], an approach that however requires even larger computational resources than RCMs. Basically, there is close connection in modelling activities of the CECILIA project to the project ENSEMBLES, where all the three models mentioned above are used as well both for climate change projections and model intercomparison and validation. As part of the ENSEMBLES project transient scenario runs of one hundred year length (and some longer) are also planned under different greenhouse gases (GHG) and aerosol forcing. In CECILIA a detailed analysis and use of the results of the project ENSEMBLES is planned for focused initial impact studies in our target region (WP1 and starting stage of WP4, 5, 6, 7). However, next to this initial phase, the main objective of the project is to adapt the RCMs used for ENSEMBLES (ALADIN-Climate and RegCM) for very high resolution (grid spacing of 10 km) simulations over selected sub-domains, which will provide additional information related to the complex terrain of the region. The assessment of the role of significant but previously not resolved topographical features and land use patterns will be provided in these experiments as well as the evaluation of the sensitivity of the simulations to the choice and size of the model domain. Further, for CECILIA project last 50 years of present century was run by the ARPEGE (CNRM) and RegCM@25km (ICTP) to provide driving fields for ALADIN-Climate and RegCM families partners, respectively, in end-of-century time slice.

In addition, there is statistical downscaling involved in the project [Busuioc et al., 2006, Huth, 2002, 2005] both for comparison and localisation of the model results. The availability of different methodologies giving often different results implies that a full assessment of the uncertainties in regional climate change simulation may require the use of multiple techniques (as proposed in this project). The CECILIA project will also provide insights on the validation and relative merits of statistical and dynamical downscaling, in particular as applied to provide local climate information.

The basic objective of modelling activities in CECILIA project is to produce simulations on targeted domains for a past period (1961-1990) driven by ERA40 reanalysis used for validation of the models as well as for a reference period (1961-1990) and scenario time slices (2021-2050 and 2071-2100) based on ENSEMBLES 6FP EC IP A1B GCM simulations. Two models have been supposed to be used as source of driving fields over six target areas, ALADIN-Climate family using stretched climate change transient run by ARPEGE/Climat for ENSEMBLES project, RegCM family using RegCM transient ENSEMBLES run for whole Europe in 25km resolution driven by transient run of ECHAM5. While the stretched ARPEGE run provides reasonable resolution in targeted regions for direct application of 10 km resolution RCM, the difference between 10 km resolution of RegCM and the resolution of other common global models is too large, that is why the double-nesting using 25 km RegCM run as an intermediate step is necessary.

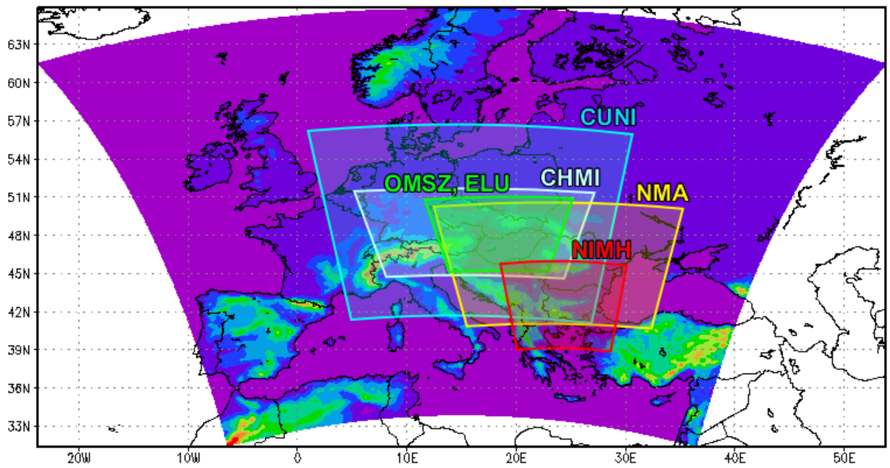


Figure 3: Integration domains for individual partner simulations

The individual partners running the simulations settled with respect to their purposes and resources available the integration domains (see Fig. 3) when preparing to perform the simulations driven by reanalysis fields, which is necessary for subsequent validation of the models performance. The first effort was given to produce the high resolution runs based on the reanalysis data in targeted areas for the period of 1961-90, allowing spinup of the models from the beginning of 1960 as starting time, moreover, most partners extended the period till the end of 2000 for more extended validation. As a source of reanalysis data, ERA 40 database has been used [Uppala et al., 2004], however, with some differences between the partners as seen in Tab. 2. Note that OMSZ and ELU share the same domain for direct results comparison. For use of ERA 40 the analysis of the impact of bigger jump between the resolutions was performed, mostly no problem considered, but improvement of the results declared at CUNI, that is why ENSEMBLES ERA 40 RegCM@25km data are used to drive CUNI simulation for reanalysis experiment. Basically, there were no significant biases with respect to high resolution CRU database reported from ALADIN – Climate family runs, but RegCM mostly giving excessive amount of precipitation – strong wet bias, especially in simulations on smaller domains. New settings of the large-scale precipitation parameterization was developed in cooperation of ELU and ICTP and tested with promising results. Exploration of the possible improvements of high resolution regional climate modelling by testing new parameterization schemes better adapted to this resolution is another objective of modelling activities within the project.

Table 2: Individual partners model and region

partner	model	resolution	domain size	domain center	boundary forcing
CUNI	RegCM	10 km	184 x 164 x 23	49.0 N, 15.8 E	ERA 40 RegCM@25
CHMI	ALADIN-Climate	10 km	148 x 72 x 43	48.30 N, 17.18 E	ERA 40
NMA	RegCM	10 km	156 x 101 x 18	46.0 N, 25.0 E	ERA 40
NIMH	ALADIN-Climate	10 km	105 x 80 x 31	42.5 N, 25.0 E	ERA 40
OMSZ	ALADIN-Climate	10 km	94 x 72 x 18	47.5 N, 18.5 E	ERA 40
ELU	RegCM	10 km	94 x 72 x 18	47.5 N, 18.5 E	ERA 40

3. FIRST RESULTS BY CUNI

In the Czech Republic a regional climate modeling program was initiated in 2001 as mentioned above using NWP model ALADIN as a basis for the development of RCM under the cooperation of CHMI, CUNI and IAP [Huth et al., 2003]. In parallel, first experiences with RegCM from ICTP were obtained at CUNI, where the analysis of RegCM ability to reproduce extremes was performed by Halenka et al. [2006]. For the CECILIA project in addition to ALADIN – Climate/CZ model developed for ENSEMBLES project and used by CHMI for domain of interest centred on Czech Republic the RegCM is run by CUNI basically over the similarly centred domain. The default setting of RegCM has been used to get extensive validation before further model settings adjustment for the region chosen. Unfortunately, the size of the domain is rather big which is prohibitive to repeat the full runs with the new settings developed.

As mentioned previously, to run the simulation driven by reanalysis the driving fields obtained from ICTP RegCM@25km run for ENSEMBLES ERA 40 experiment were used instead of direct use of ERA40 fields. The preliminary tests using the first six years of simulation analysed vs. station data show the better performance of the model in such a case, moreover, for further scenarios analysis the advantage is that all these simulations will be performed exactly the same way.

Comparison of reanalysis run for temperature with respect to CRU 10‘ climatology is was performed for January and July, respectively, in terms of individual decades. Reasonable agreement can be seen, the similar patterns in interdecadal variations, like the coldest first decade in January, the warmest third decade and colder second one for July temperature can be seen in the simulation. The benefit of high resolution is well expressed especially when analysing local scale (not presented). The biases for temperature both in January and July are shown in Fig. 4, there is slight warm bias in January in most region of interest, with pronounced tendency to none or colder bias in the upper locations which might be another demonstration of high resolution benefit as orography of the model terrain is better represented than orography in 10' CRU resolution. This is similar to July performance where rather cold bias prevails anyway. Overall performance for temperature thus looks reasonably good.

However, this is not the case for precipitation fields. There is rather overestimation of precipitation due to excessive smaller amounts appearance. Significant wet bias can be seen both for January and July in Fig. 5. The problem is not so serious in main domain of interest during July (except the first decade), for January there is again the feature of upper locations expressed, but further analysis has to be performed to get sources of the differences, as well as the tests with modified precipitation parameterization settings are necessary. Similarly as for temperature, despite of the biases the potential of high resolution simulation in capturing of orographic features is well pronounced.

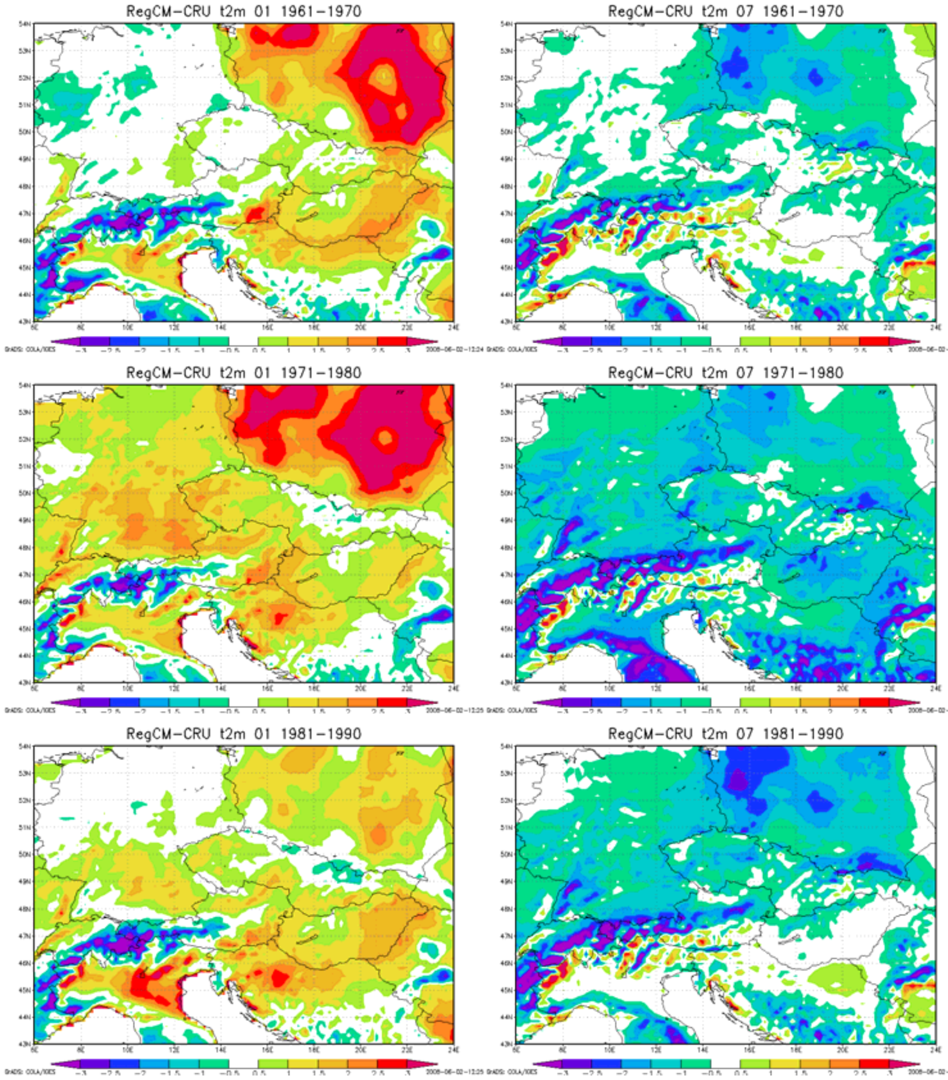


Figure 4: Temperature in 2 m, decadal average difference of simulations against CRU@10'. January on the left panels, July on the right panels.

The density of information in high resolution and similarity of model terrain and the real topography enables the application of further localization technique adopted by CUNI. It is based on regression analysis of the dependence of the climate characteristics on the orography height in the vicinity of the point of interest, usually, for validation purposes, the observational site, but for an analysis of scenarios runs e.g. the place for impact study. Thus it is possible to get the value of the parameter on real topography and to go farer beyond the 10km of model simulation resolution (see Fig. 6). It should be mentioned that these relations are quite realistically represented in the models especially for temperature parameters where this is more straightforward, but e.g. for ALADIN – Climate it is surprisingly well represented for precipitation as well. There is one advantage of this technique, as it is based on the analysis of model data only, it can be used the same way for scenarios runs even reproducing possible changes in these dependencies.

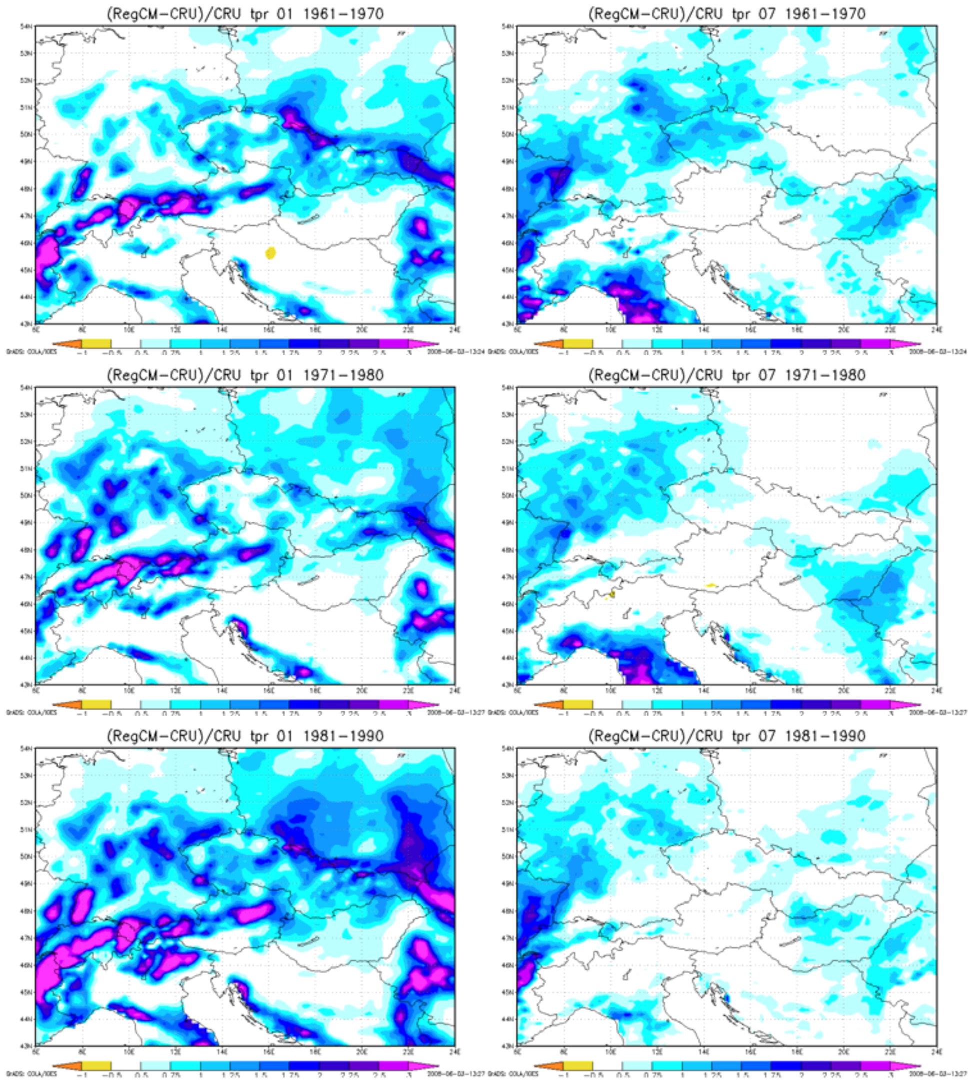


Figure 5: Monthly precipitation, decadal average relative bias of simulations against CRU@10'. January on the left panels, July on the right panels.

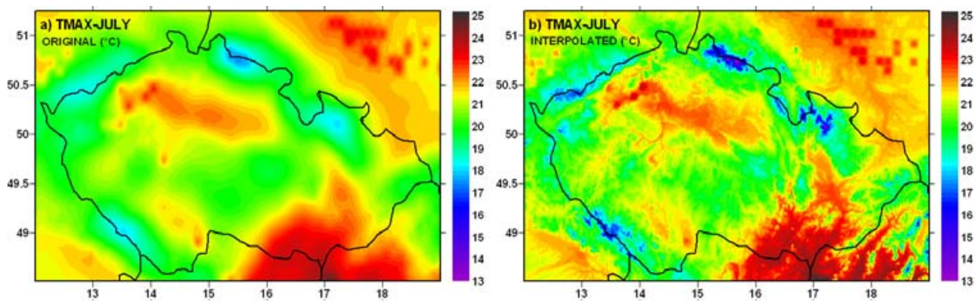


Figure 6. Monthly maximum temperature in July, simulated by the RegCM model (left panel) and after the altitude-based localization of the model outputs (right panel).

4. CONCLUSIONS

The aim of the project is to assess the impact of climate change at the regional to local scale for CEE using very high resolution simulations in order to include the effects of the complex terrain of the region. This is achieved mainly using very high resolution RCMs run locally on rather limited areas for individual targeted areas which enables the simulations even with quite limited resources, eventhough the amount of produced data is enormous. From the viewpoint of climate change scenario production two time slices are planned and in a few cases achieved, for 2020-2050 and 2070-2100. The selected applications of the CECILIA outputs are dealing with water resources and management, agriculture, forestry, air quality and health. In addition, CECILIA is improving the access of CEE researchers to information and facilities for climate change research by providing an efficient use and access to the results of previous and ongoing EC projects which the proposed research benefits greatly from, e.g. Modelling the Impact of Climate Extremes (MICE), Statistical and regional dynamical downscaling of extremes for European regions (STARDEX), Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects (PRUDENCE) and ENSEMBLE-based Predictions of Climate Changes and their Impacts" (ENSEMBLES).

The most reliable source of information on the evolution of the atmospheric environment in the next decades comes from RCMs. The project CECILIA brings for the first time very high resolution localization of climate change scenarios into the targeted areas of CEE. Improving upon the project ENSEMBLES where Europe-wide scale is adopted at high resolution, here we address even higher resolutions on a significantly smaller domain. This higher resolution enables not only more detailed description of the topography and land use, but it allows to introduce new processes, as interactive interaction of climate change and air quality, subgrid effects etc. The preliminary results show the promising benefits of very high resolution towards the use in impact studies. More detailed analysis of results which is clearly beyond the purpose of this paper is either already done or it is planned in framework of the project. Moreover, development of new features in the parameterization of high resolution physics in the models is expected. Adjustment of large scale precipitation parameterization has been mentioned already as improving the precipitation results significantly in very high resolution. WUT has studied an alternative boundary layer scheme in RegCM, subscale effect parameterization in RegCM is tested in Alpine region by BOKU etc. In addition, in framework of impact studies on air quality the coupling of RegCM with chemistry transport model is used, studies analysing the climate change impact on different atmospheric processes and interactions with the Earth surface are performed as well. The CECILIA project will also provide insights on the validation and relative merits of statistical and dynamical downscaling, in particular as applied to provide fine scale climate information. Previous results show the possibility of the changes of statistical distribution of climate parameters in our targeted domains. Despite the relative agreement of climate-change scenarios concerning the changes in extremes over this region, a significant uncertainty remains with regard to their exact magnitude and the attribution of the causes for these changes. Some studies have highlighted the role of large-scale circulation changes, but land-atmosphere interactions are clearly of key relevance as well. Moreover, certain aspects central to this issue are often not well represented in GCMs (land surface heterogeneity, complex topography, convection), or even in coarser RCMs. Very high-resolution simulations seem to help significantly investigating some of these open questions and yield more accurate estimates of future changes in extreme weather events over the targeted regions.

ACKNOWLEDGEMENTS

This work is supported in framework of EC FP6 STREP CECILIA (GOCE 037005). Some contributions to the tasks are supported from local sources as well under Research Plan of MSMT, No. MSM 0021620860, and in framework of the project supported by Czech Science Foundation under No. 205/06/P181. Authors wish to express their thanks to all CECILIA team, especially the modelling partners, we acknowledge their contributions to the overall review of modelling activities within CECILIA project.

REFERENCES

Beniston, M., and D.B. Stephenson: Extreme climatic events and their evolution under changing climatic conditions. *Global and Planetary Change*, 44, 1-9, 2004.

Bubnová, R., Hello G., Bénard P. and Geleyn J.F.: Integration of the fully compressible equations with orography in the framework of the ARPEGE/ALADIN NWP. *Meteorol. zpr.*, 47, 69-76, 1994.

Busuioc, A, F. Giorgi, X. Bi and M. Ionita: Comparison of regional climate model and statistical downscaling simulations of different winter precipitation change scenarios over Romania. *Theor. Appl. Climatol.*, 86, 1-4, 101-123, 2006.

Déqué, M, Piedelievre, J.P.: High-Resolution climate simulation over Europe. *Clim. Dyn.*, 11, 321-339, 1995.

Déqué, M, Marquet, P., Jones, R.G.: Simulation of climate change over Europe using a global variable resolution general circulation model. *Clim. Dyn.*, 14, 173-189, 1998.

Dickinson, R., A. Henderson-Sellers, and P. J. Kennedy: Biosphere-Atmosphere Transfer Scheme (BATS) version 1E as coupled to the NCAR Community Climate Model. NCAR Tech. rep. TN-387+STR, 72 pp., 1993.

Elguindi, N., X. Bi, F. Giorgi, B. Nagarajan, J. Pal, F. Solmon, S. Rauscher, A. Zakey: RegCM Version 3.1 User's Guide. PWCG Abdus Salam ICTP, 2006.

Frei, C., J.H. Christensen, M. Deque, D. Jacob, R.G. Jones, and P.L. Vidale: Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps. *Journal of Geophysical Research*, 108 (D3), 4124, 2003.

Gates, W. L., A. Henderson-Sellers, G. J. Boer, C. K. Folland, A. Kitoh, B. J. McAvaney, F. Semazzi, N. Smith, A. J. Weaver, and Q.-C. Zeng: Climate models - Evaluation, in Climate Change 1995: The Science of Climate Change, edited by J. T. Houghton et al., pp. 229–284, Cambridge Univ. Press, New York, 1996.

Giorgi, F., M.R. Marinucci, and G.T. Bates: Development of a second generation regional climate model (RegCM2). Part I: Boundary layer and radiative transfer processes. *Mon. Wea. Rev.*, 121, 2794-2813, 1993a.

Giorgi, F., M.R. Marinucci, G.T. Bates, and G. DeCanio: Development of a second generation regional climate model (RegCM2). Part II: Convective processes and assimilation of lateral boundary conditions. *Mon. Wea. Rev.*, 121, 2814-2832, 1993b.

Giorgi F, L.O. Mearns: Introduction to special section: regional climate modeling revisited. *Journal of Geophysical Research*, 104, 6335–6352, 1999.

Giorgi, F., Y. Huang, K. Nishizawa and C. Fu: A seasonal cycle simulation over eastern Asia and its sensitivity to radiative transfer and surface processes. *Journal of Geophysical Research*, 104, 6403-6423, 1999.

Giorgi, F., X. Bi and J. Pal: Means, trends and interannual variability in a regional climate change experiment over Europe. Part I: Present day climate (1961-1990). *Climate Dynamics*, 22, 736-756, 2004.

Giorgi F., E. Coppola: European climate-change oscillation (ECO), *Geophys. Res. Lett.*, 34, L21703, doi:10.1029/2007GL031223, 2007.

Giorgi, F., Francisco, R.: Evaluating uncertainties in the prediction of regional climate change. *Geophys. Res. Lett.*, 27, 1295-1298, 2000.

Giorgi, F., Hewitson, B., Christensen, J., Hulme, M., von Storch, H., Whetton, P., Jones, R., Mearns, L., Fu, C.: Regional climate information – evaluation and projections. In: Climate Change 2001: The Scientific Basis, Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., van der Linden, P.J., Dai, X., Maskell, K., Johnson, C.A., Eds. Cambridge University Press, Cambridge, 583-638, 2001.

Giorgi, F., Y. Huang, K. Nishizawa and C. Fu: A seasonal cycle simulation over eastern Asia and its sensitivity to radiative transfer and surface processes. *Journal of Geophysical Research*, 104, 6403-6423, 1999.

Gobron, N., B. Pinty, F. Melin, M. Taberner, M.M. Verstraete, A. Belward, T. Lavergne and J.L.Widlowski: The state of vegetation in Europe following the 2003 drought. *Int. J. Remote Sens.*, 26, 2013-2020, 2005.

Hafner, S.: Trends in maize, rice and wheat yields for 188 nations over the past 40 years: a prevalence of linear growth. *Agr. Ecosyst. Environ.*, 97, 275-283, 2003.

Halenka, T., J. Kalvová, Z. Chládová, A. Demeterová, K. Zemánková, and M. Belda: On the capability of RegCM to capture extremes in long term regional climate simulation – comparison with the observations for Czech Republic, *Theor. Appl. Climatol.*, 86, 1-4, 125-145, 2006.

Huntingford, C., R.G. Jones, C. Prudhomme, R. Lamb, J.H.C. Gash, and D.A. Jones: Regional climate-model predictions of extreme rainfall for a changing climate. *Quart. J. of Royal Met. Soc.*, 129, 1607-1621, 2003.

Huth, R.: Statistical downscaling of daily temperature in central Europe. *J. Climate*, 15, 1731-1742, 2002.

Huth, R., J. Kyselý, and M. Dubrovský: Simulation of surface air temperature by GCMs, statistical downscaling and weather generator: higher-order statistical moments. *Studia geoph. geod.*, 47, 203-216, 2003.

Huth, R., R. Mládek, L. Metelka, P. Sedlák, Z. Huthová, S. Kliegrová, J. Kyselý, L. Pokorná, M. Janoušek, T. Halenka: On the integrability of limited-area numerical weather prediction model ALADIN over extended time periods. *Studia geoph. geod.*, 47, 863-873, 2003.

Huth, R.: Downscaling of humidity variables: A search for suitable predictors and predictands. *Int. J. Climatol.*, 25, 243-250, 2005.

Kysely, J.: Mortality and displacedmortality during heat waves in the Czech Republic. *Int. J. Biometeorol.*, 49, 91-97, 2004.

Meehl, G.A., C. Tebaldi, and D. Nychka: Changes in frost days in simulations of 21st century climate. *Clim. Dyn.*, 23, 495-511, 2004.

Meehl, G.A., and C. Tebaldi: More intense, more frequent and longer lasting heat waves in the 21st century. *Science*, 305, 994-997, 2004.

Menzel, A., G. Jakobi, R.Ahas, H. Scheifinger and N. Estrella: Variations of the climatological growing season (1951-2000) in Germany compared with other countries. *Int. J. Climatol.*, 23, 793-812, 2003.

Menzel, A., T.H. Sparks, N. Estrella, E. Koch, A.Aasa, R.Ahas, K. Alm-Kubler, P. Bissolli, O. Braslavska, A. Briede, F.M. Chmielewski, Z. Crepinsek, Y. Curnel, A. Dahl, C. Defila, A. Donnelly, Y. Filella, K. Jatczak, F. Mage, A. Mestre, O. Nordli, J. Penuelas, P. Pirinen, V. Remišova, H. Scheifinger, M. Striz, A. Susnik, A.J.H. vanVliet, F.-E. Wielgolaski, S. Zach and A. Zust: European phonological response to climate change matches the warming pattern. *Glob. Change Biol.*, 12, 1969-1976, 2006b.

Pal, J. S., E. E. Small, and E. A. Eltahir: Simulation of regional-scale water and energy budgets: Representation of subgrid cloud and precipitation processes within RegCM. *J. Geophys. Res.*, 105, 29579-29594, 2000.

Pal, J. S., F. Giorgi, and X. Bi: Consistency of recent European summer precipitation trends and extremes with future regional climate projections. *Geophysical Research Letters*, 31, L13202, doi:10.1029/2004GL019836, 2004.

Pal, J.S., F. Giorgi, X. Bi, N. Elguindi, F. Solmon, X. Gao, S.A. Rauscher, R. Francisco, A. Zakey, J. Winter, M. Ashfaq, F.S. Syed, J.L. Bell, N.S. Diffenbaugh, J. Karmacharya, A. Konaré, D. Martinez, R.P. da Rocha, L.C. Sloan, and A.L. Steiner: Regional Climate Modeling for the Developing World: The ICTP RegCM3 and RegCNET. *Bull. Amer. Meteor. Soc.*, 88, 1395–1409, 2007.

Schär, C., P.L. Vidale, D. Luthi, C. Frei, C. Haberli, M.A. Liniger, and C. Appenzeller: The role of increasing temperature variability in European summer heat waves. *Nature*, 427, 332-336, 2004.

Uppala, S., Kallberg, P., Hernandez, A., Saarinen, S., Fiorino, M., Li, X., Onogi, K., Sokka, N., Andrae, U., Da Costa Bechtold, V.: ERA-40: ECMWF 45-year reanalysis of the global atmosphere and surface conditions 1957-2002. *ECMWF Newsletter*, 101, 2-21, 2004.

Váňa F.: Physical parametrizations in the ALADIN model. *Meteorol. zpr.*, 51, 33-44, 1998. [in Czech]

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Rapid deployment system for complex accidental release simulations

P. Juruš[a], K. Eben[a], J. Resler[a] and M. Belda[b]

[a] *Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 2, 182 07 Prague 8, Czech Republic (jurus@cs.cas.cz)*

[b] *Faculty of Mathematics and Physics, Charles University, V Holešovičkách 2, 180 00 Prague 8, Czech Republic*

Abstract: State-of-the-art numerical models of weather and air quality can be used to model certain types of accidental atmospheric releases. This can lead to much improved forecast compared to simple Gaussian dispersion models which are traditionally used in the domain of atmospheric release modelling. To setup and run such a sophisticated model is however task which requires considerable effort and time, hindering the use of complex models in accidental release modelling and risk management planning. To overcome these problems, we have developed ICT system ARRAS (Atmospheric Release Risk Assessment System). By using the ARRAS, the complex numerical models of atmosphere can be configured for accidental release simulation by filling the simple web-based data form. System afterwards automatically executes the numerical model and when the first results are available, it presents them on the web page. User-friendly interface thus greatly speeds up the time needed for the simulation setup, and broadens the family of models that can be used for accidental release simulation.

Keywords: Environmental modelling; Accidental release modelling; Risk management planning.

1. INTRODUCTION

Release of hazardous materials to the atmosphere either by an accident or even due to terrorist action belongs to the serious threats of the modern civilization [Ebel, Davitashvili, 2007]. Instruments for the rapid response are essential in such a case. The dispersion models together with detailed weather prediction are important tools for the real-time nowcasting and prediction of exposure and health impact analysis. These models are used also for case studies aimed at accidental release prevention and mitigation as a part of risk management planning.

Traditionally simple Gaussian models are used in such settings. Their advantages are robustness, low computational requirements as well as relative simplicity of use and ability to give results even with very limited weather data. These qualities make such models particularly suitable shortly after the accident when the information about the release is often limited and the rapid estimate of further dispersion is needed. The disadvantages of traditional Gaussian models are the simplifications, which include weather description, dispersion and eventual deposition. These models in most cases completely bypass atmospheric chemistry.

State-of-the-art numerical models of weather and air quality are able to model the atmosphere on regional scale. Higher computational demands are compensated by the much more detailed physical description of reality. This makes possible to model chemical species dispersion with the inclusion of e.g. vertical diffusion, precipitation, deposition, or chemical reactions. These more complex models are therefore useful especially for the larger scale events, where the accurate prediction is needed for timescales ranging from hours to several days and the area of interest is of the regional scale (from hundreds of meters to hundreds of kilometres).

The complexity of numerical models of atmosphere is however one of the biggest obstacles that hinders their usage in risk management systems. The ICT system ARRAS which is being developed in the Institute of Computer Science in Prague is designed to close the gap between the state-of-the-art numerical models of atmosphere and accidental release risk management systems.

2. ARRAS DESIGN

The main design goals of the system are:

1. Straightaway interface for the submission of the information about the accidental release.
2. Fully automated configuration of numerical models of atmosphere, dynamically adapted according to accidental release information and to the desired area of interest.
3. Runtime information about the actual state of computation and the estimate of completion time.
4. Comprehensive information about the predicted weather and dangerous species concentrations.
5. Continuous update of the available results while the models are still running.

The components of the system are following:

1. System core using the numerical weather prediction model WRF.
2. Chemical transport model coupled to the weather outputs.
3. Postprocessing components used for the data processing into the useful formats (e.g. the graphical maps of predicted concentrations).
4. Web-based user interface for both submitting the accidental release information and executing the simulation, and presenting the simulation outputs.

The system uses the WRF (Weather Research & Forecasting) model for the numerical weather prediction [Skamarock et al., 2005]. It is fully nonhydrostatical model and allows weather simulation in fine horizontal scales. Simulation domain is automatically created for each run so that site of accidental release is at its centre. The spatial resolution is 1km and 33 vertical levels is used. Model uses so-called NOAH land-surface model with 4 soil layers, and standard widely-used physical options and parameterisations. The time step of numerical integration is 3s. Outputs given by WRF are 3-dimensional fields of the meteorological variables in NetCDF format. The time step of output is 1 hour with the possibility of finer time resolution in case of need.

The availability of operational mesoscale numerical weather forecasting system is essential for the operation of ARRAS. We use the MEDARD system [Eben et al., 2005] for this purpose and it serves as the source of initial and boundary conditions for the ARRAS. The system MEDARD is based on the WRF model as well and uses nested domains with horizontal resolution 27, 9 and 3km.

Components for the coupling with chemical transport model depend on the chemical model used. In principle any model, which can use subset of meteorological outputs provided by WRF, can be used. This means that number of Gaussian dispersion models could be used for example. Moreover such a model can be driven not only by the predicted or measured wind at the site of accident, but also by the predicted and time-varying wind field in the whole simulated domain (for the possible effects see Figure 1.). This can bring substantial improvement of the forecast, especially for the longer time horizons.

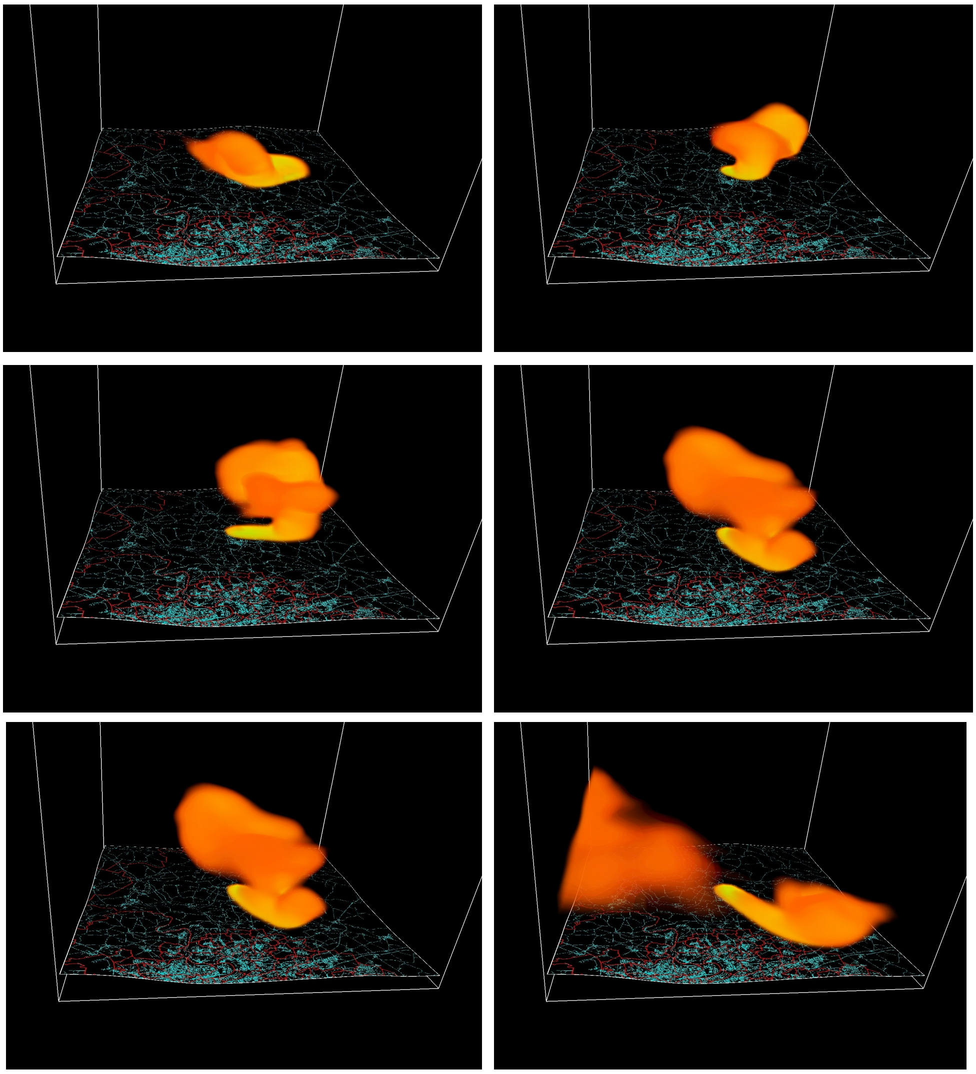


Figure 1. 3-dimensional visualisation of the accidental release shows the effects of differing wind direction in higher levels above ground. The toxic cloud is divided into the two parts each travelling into the different location. This effect of dynamically changing wind field couldn't be modelled by the majority of simpler models. The visualisation was performed by the VAPOR – visualisation and analysis platform [Clyne et al., 2007].

3. IMPORTANT FEATURES OF ARRAS

Here is the summary of the most important features of the system ARRAS that differentiate it from the most of widely used systems for the accidental release modelling.

- The core of the system is numerical weather prediction model (WRF in the case of the prototype).
- System uses outputs of operational numerical weather forecast (system MEDARD www.medard-online.cz in the case of the prototype).
- Ability to use the most precise weather forecast available for the simulated incident.
- Modularity - ability to use different models for chemical transport. (CMAQ and WRF-Chem used in the prototype)

Our intention was to demonstrate to possibility of coupling with sophisticated Eulerian models. Such models can for example simulate chemical reactions or wet deposition by precipitation, in case that specific accidental release conditions suggest that a more complex simulation besides the simple dispersion is needed. Fully 3-dimensional flow can also have significant effect, especially when the wind speed and direction differs a lot in various vertical layers. We have tested the ARRAS prototype with the CMAQ model and with WRF extension WRF-Chem [Grell et al., 2005]. In the latter case WRF-Chem serves as both numerical weather prediction model and chemical transport model.

4. USER INTERFACE

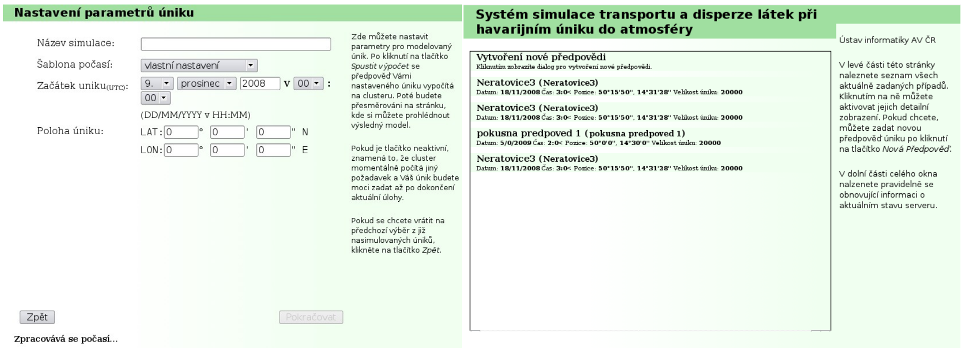


Figure 2. Sample of the user interface. Form for submitting the accidental release information is on the left side, list of browsable already computed scenarios is on the right. This particular interface is localized into the Czech language.

User interface is designed to be as simple as possible. When the accidental release occurs, the end-user only submits the basic information like the coordinates and the time of release. Chain of fully automated operations is started afterwards and whole process ends by publishing of the results in the graphical form on the web.

Model domain is created at the beginning. This domain contains 40x40 gridpoints with the centre at the accidental release site and the horizontal resolution of the grid is 1km. Preparation of the boundary and initial conditions is the second step. The MEDARD outputs are used in the prototype. Fine scale weather prediction starts after the interpolation of the boundary and initial conditions for the given domain. The outputs of numerical weather prediction can be visualised by the RIP (Read, Interpolate, Plot – part of the WRF distribution) postprocessor.

End user can submit more detailed parameters of the accidental release during the computation of the weather. User can either define the parameters from the scratch or use one of the predefined cases or previous simulations that are offered by the web interface. The chemical transport model can be run as soon as the numerical weather model finishes. System checks the existence of the outputs and displays the results consecutively as soon as they are available. The simulation ends when the output for the desired forecast horizon (specified in the system configuration) is published.

Special flag prevents the execution of another simulation while there is already another simulation running. End user (or potentially even more users at different places) can however at any time study already finished simulation and monitor the running simulation because the actual simulation state is reported and expected time of the results publication is displayed. The web interface can be used for the interactive analysis of the previous simulations. System currently supports one numerical weather forecast for each episode but potentially more runs of chemical transport model. Different runs can use different emission parameters. This makes possible creation of the different scenarios, helping to analyse the uncertainty of the forecast and sensitivity to emission size.

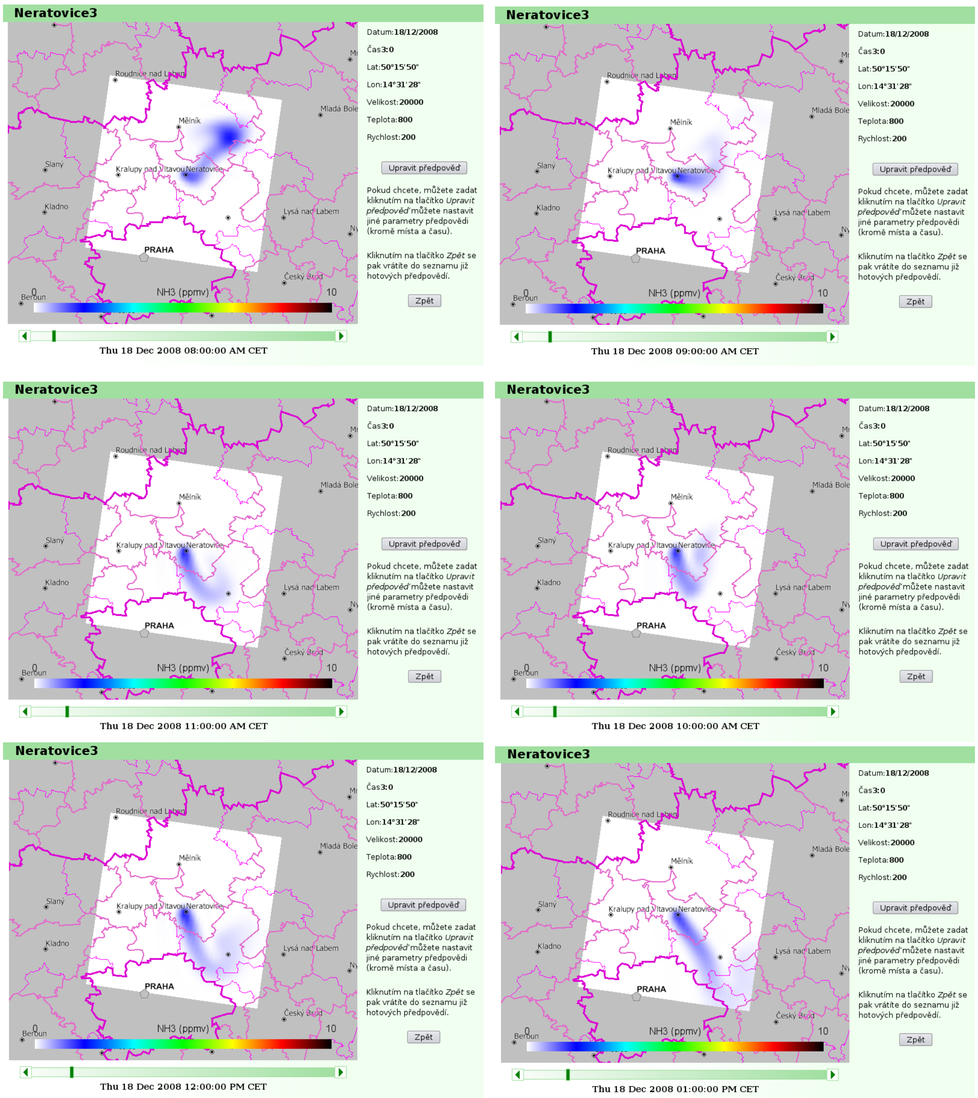


Figure 3. An example of the ARRAS output, simulated accidental release in the vicinity of Prague.

5. POSSIBLE USES OF THE SYSTEM

Computational time is one of the most important factors for the real-life utilisation of the system. We have tested the prototype in the Institute of Computer Science in Prague using two workstations. Each workstation has four quad core processors giving total of 32 processor cores used for the model runs. The workstations are interconnected by the Infiniband network and thanks to well parallelized code of the numerical models, the computational time decreases almost in reciprocal proportion to the number of processor cores. In the tested configuration 1 hour of simulation took approximately 1 minute of computation time.

The ARRAS system is modular, highly customizable and can be used in different settings and in different ways. Here are some typical examples:

- Single purpose operational application for the modelling of one or more types of dangerous species for the specific site. Profiles of potential accidents can be preconfigured in the system.
- Interactive analysis of different accidental release scenarios for different meteorological situations. Support tool for creation of risk management plans for the particular site.

- Utilisation of the meteorological part of the system alone to get the forecast with fine horizontal resolution.

6. CONCLUSIONS

The new ICT system ARRAS for the accidental release modelling has been developed in the Institute of Computer Science in Prague. One of the most distinctive features of the system is utilisation of full fine-scale numerical weather forecast and state-of-the-art chemistry transport model. Another important feature is user friendly web-based interface as well as modular design allowing easy change of the numerical models and further customization. The prototype of the system indicates that the system could be especially useful for the modelling of the larger scale accidents and in more complex weather situations.

ACKNOWLEDGEMENTS

This work was supported by the Grant Agency of the Academy of Sciences of the Czech Republic (grant No. 1ET400300414, framework 'Information Society') and by the Institutional research plan AV0Z 10300504 Computer Science for the Information Society 'Models, Algorithms, Applications'.

REFERENCES

Ebel, A., and T. Davitashvili (eds.), Air, Water and Soil Quality Modelling for Risk and Impact Assessment, NATO Science for Peace and Security Series C: Environmental Security, Springer, 2007.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang and J. G. Powers, A Description of the Advanced Research WRF Version 2, NCAR Technical Note, NCAR/TN-468+STR., 2005.

Grell, G. A., S. E. Peckham, R. Schmitz, S. A. McKeen, G. Frost, W. C. Skamarock and B. Eder, Fully coupled online chemistry within the WRF model, *Atmospheric Environment*, 39, 6957-6975, 2005.

Eben, K., P. Juruš, J. Resler, M. Belda, B. C. Krüger and J. Keder, MEDARD – an Environmental Modelling Project for the Territory of the Czech Republic, *ERCIM News*, 61, 2005.

Clyne, J., P. Mininni, A. Norton and M. Rast, Interactive desktop analysis of high resolution simulations: application to turbulent plume dynamics and current sheet formation, *New Journal of Physics,* 9, 301, 2007.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

REACh – New IT-Trend in collaboration and networking for the Process of Registration, Evaluation, Authorisation and Restriction of European Chemicals

G. Knetsch[a]

[a] *Federal Environment Agency,06844 Dessau-Roßlau,Wörlitzer Platz 1, Germany (Gerlinde.Knetsch@uba.de)*

Abstract: REACh is the regulation for Registration, Evaluation, Authorisation and Restriction of Chemicals. It entered into force on 1st June 2007 to streamline and improve the former legislative framework on chemicals of the European Union. REACh places greater responsibility on industry to manage the risks that chemicals may pose to the health and the environment. In principle REACh applies to all chemicals, not only chemicals used in industrial processes but also in our day-to-day life, for example in cleaning products, paints as well as in articles such as clothes, furniture and electrical appliances. The concept of REACh in a short term means: No data – no market! The application of the REACh processes requires the use of several tools or methodologies, existing or developed for the purpose of REACh. To ensure support to all REACh processes, Information Technology tools have been developed to submit, store and exchange information and data on chemicals. REACh-IT is the portal for the management of the workflows, including the pre-registration process, dossier submission, as well as the data base and application of IUCLID 5 and the website of the European Chemical Agency.

Keywords: Chemicals; Registration; REACh; Workflow; WebService;

1. INTRODUCTION

In December 2006, the Regulation No 1907/2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACh) and establishing a European Chemicals of the EU Agency (ECHA) was adopted. The objectives of the regulation are:

- To improve the protection of human health and the environment from the risks that can be posed by chemicals,
- To enhance the competitiveness chemicals industry, a key sector for the economy of the EU,
- To promote alternative methods for the assessment of hazards of substances,
- To ensure the free circulation of substances on the internal market of the European Union.

Under the Directive 67/548/EEC and actions taken by The European Chemical Agency (ECHA) under Article 16(1) for the inclusion in the database of information of substances regarded as registered under Article 15 a complex and extensive IT-network and collaboration system has been installed. This system allows managing the pre - registration, dossier submissions (registration, PPORD[1] notification and inquiry) and assignment of registration numbers for notified substances. REACh-IT ensures an efficient, consistent and secure exchange of data and information for the evaluation, authorisation and restriction processes for chemical substances across the European Union. The REACh-IT network is designed to provide additional information on chemicals, to ensure their safe use and to ensure competitiveness of the European industry. It also allows the Agency and Member States authorities to download dossiers for review processes via and VPN2 connection. The Agency will also use REACh-IT to make non-confidential information on chemicals accessible to public on its website.

2. OBJECTIVES AND USER GROUPS

The aim of the REACh-IT project is to develop IT-support for the implementation of the REACh legislation and the daily work of the Agency (ECHA) and the Member States Competent Authorities (MSCAs), enabling them, together with Industry, to fulfil their duties and obligations as defined in the REACh legislation. The main objectives of REACh-IT are:

- To develop an IT-system that enables the ECHA to receive and store pre-registration and inquiry dossiers, registrations, notifications, reports and payments
- To make the information easily retrievable by the Agency and MSCAs
- To enable the dissemination of non-confidential registration dossier data and other relevant data.

Three main types of actors in REACh-IT are in the focus and the retrieval of the data has to be profile specific. The user groups are the Industry (manufacturers, importers, representatives, downstream user incl. distributors), the Agency / MSCAs (Agency staff, MSCA staff) and the general public (consumer and environmental NGOs).

It is very important to define roles and rules for the processes for data submitting and managing. The profile specific retrieval is one of the aspects for the security network of REACh-IT. Figure 1 reflects the three pillars of the user groups.

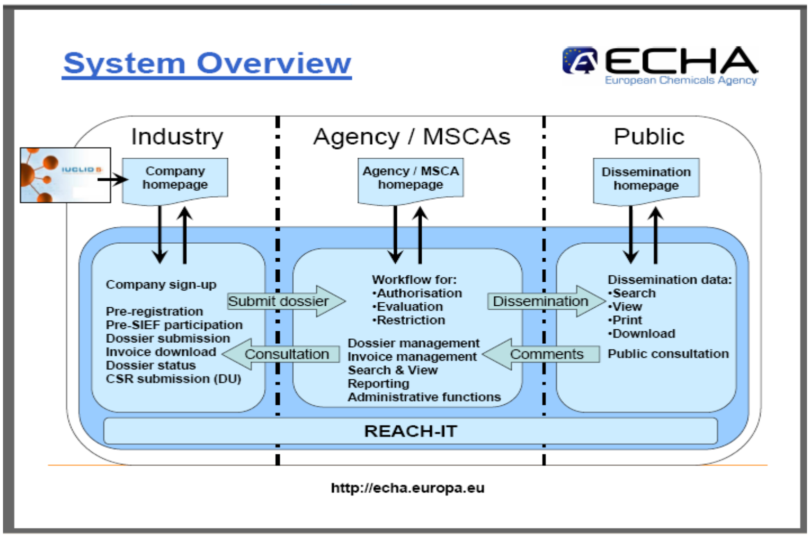


Figure 1: Overview of the three user groups of REACh-IT
Source: ECHA 2008a

[1] PPORD: **P**roduct and **P**rocess Oriented **R**esearch and **D**evelopment
[2] VPN: Virtuell Private Network

The development of XML-interfaces supplies these functions and gives the user groups the opportunity to login with a special profile. Another aspect is the down- and uploading function of data for authorised users. There are some rules:

- Each type of actor has its specific needs.
- Each type of actor has its specific security requirements.
- Separate web-sites for each actor.
- Agency and MSCAs have access to the same type of functionality.

3. IT -TOOLS SUPPORTING THE PROCESSES OF REACH

3.1 Formats and Software for submission

The REACh Article 111 "Formats and software for submission of information to the Agency" states, "The Agency shall specify formats and make them available free of charge [...] on its website for any submissions to the Agency. Member States, manufactures, importers, distributors or downstream users shall use these formats [...] in their submissions to the Agency pursuant to this Regulation. [...] For the purposes of registration, the format of the technical dossier [...] shall be IUCLID 5. The Agency shall coordinate the further development of this format with the Organisation for Economic Cooperation and Development to ensure maximum harmonisation."

To fulfil this obligation the ECHA has been developed IUCLID 5 - the first IT system to implement completely the so-called OECD Harmonised Templates. In context of the OECD project on Harmonised Templates for reporting summary information from testing results on chemical safety, 87 templates were developed and published in March 2006. These templates are the backbone of the IT-application of IUCLID 5.

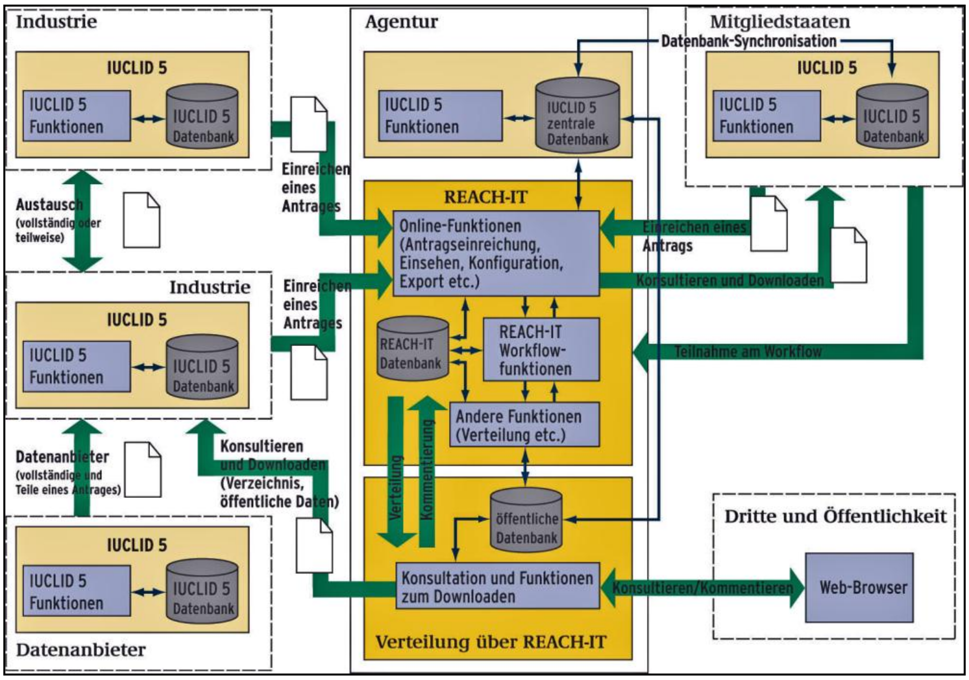


Figure 2. REACh-IT and IUCLID 5 – Workflow and Management System, Source: ECHA, modified by UBA (2007)

3.2 IT-Tools for Pre-Registration Process

During the pre-registration period of six months - from the 1th of June 2008 until 31th November 2008 – more than 65.000 companies have signed up REACh-IT and submitted more than 2.6000.000 pre-registrations. The ECHA offered three possibilities:

- manually entering online the required data in REACH-IT,
- using the pre-registration plug-in of IUCLID 5 to generate an XML file and import it to REACH-IT for submission,
- using own company IT system to create an XML file and import it to REACH-IT for submission.

The XML interface in IUCLID 5 is the key for harmonising and exchanging file, on one hand for pre-registration and on the other hand for submitting dossiers. In figure 3 the task panel shows four functionality groups:

- Tasks for creation Legal entities and Legal entity site, Substance, Mixture, Template, Category and View Dossier data
- Inventories
- Tools and administration
- PlugIns for Query tool, Migration SNIF data and IUCLID 4 data

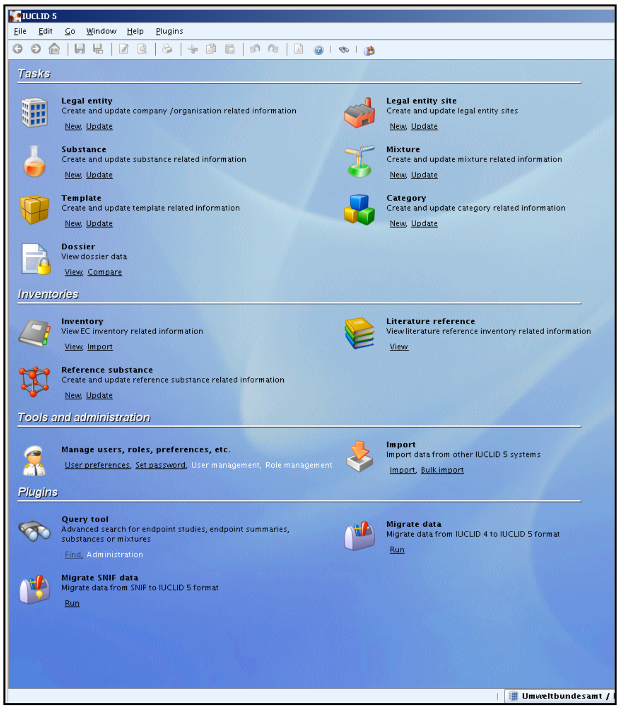


Figure 3. Screenshot of the IUCLID 5 application with four functional groups

3.3 IT-Tools for extended query possibilities

The IUCLID 5 Query Tool – available as a plugin since February 2009 - provides easy access to predefined sets of business-relevant information. It can be used to perform user-defined queries on both the administrative (e.g. company information, tonnage band, confidentiality flags) and the scientific data fields (e.g. substance datasets, materials and methods, study results). All dossiers are stored in the IUCLID 5 database. Hundreds of fields can be queried to retrieve only the relevant information from endpoint and study summaries recorded in substances, mixtures or templates. A wide range of query combinations are possible with a user friendly interface. Up to 10 query blocks can be defined to retrieve raw data, dossier data or both in combination. An overview about the query block allows selecting the block for witch to define search criterias: Substance, Mixture, Templates, Annotations, Dossier Headers, Endpoint Studies and Endpoint Summaries

Results for the query are structured in a well formed way. Functionalities are implemented to link the short overview with more detail information about the item e.g. physical-chemical properties.

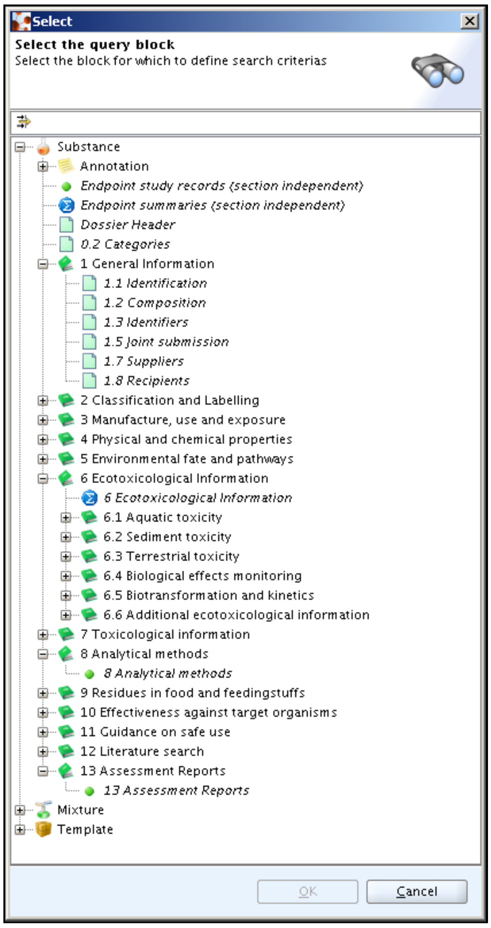


Figure 4. Screenshot of the IUCLID 5 Query blocks and search criterias

3.4 IT-Tools for preparing Chemical Safety Assessment

In 2007/2008 the European Commission, in co-operation with the ECHA has carried out an Analysis and Design Study to provide industry with IT-tools to generate **E**xposure **S**cenarios (ES), **C**hemical **S**afety **A**ssessments (CSA) and **C**hemical **S**afety **R**eports (CSR). For manufacturers and importers of substances it is very important to:

- gather information on the environmental and health properties of their substances,
- assess the risks arising from the uses of their substances and
- ensure that these risks are properly managed.

To demonstrate these they have to submit:

- technical dossier, for substances in quantities of 1 tonne or more per year, and, in addition,
- **C**hemical **S**afety **R**eport, for substances in quantities of 10 tonnes or more per year.

The **C**hemical **S**afety **R**eport documents the **C**hemical **S**afety **A**ssessment consisting of:

- the hazards assessment,
- classification of the substance and
- the assessment whether the substance is persistent, bio-accumulative and toxic (PBT), or very persistent and very bio-accumulative (vPvB).

Parts of information for these items have to be stored in IUCLID 5, containing in the dossiers. It is foreseen to implement an IUCLID 5 plug-in with following functionalities: Collect all data out of the IUCLID 5 data structure for a single substance, which are relevant for the CSR. Then there will be an interface for generating a well structured document in a standard format out of the collected data of IUCLID 5. This document can be completed and updated manually by the user.

It is foreseen, that the ECHA's CSA Tool in its first version provides a workflow support for registrant in carrying out the Assessment for identified uses, to conduct exposure estimates and to generate an Exposure Scenario and a Chemical Safety Report in a standard format.

5. CONCLUSIONS

The complex REACh-IT system is a project between the European Chemical Agency, the Member States of Europe and the Industry, financed by the European Commission. The development of the tools and the technical infrastructure has been highly influenced by the unique requirements of the several user groups. The discussion of possible approaches in the context of the applications REACh-IT and IUCLID 5 has been closely related to security aspects for confidential data. The submission of dossiers needs a special technical infrastructure with cryptoboxes and a security network so called VPN.

The REACh-IT Portal and the IUCLID 5 application are developed under the preconditions of the W3C-Standards and of Open-Source Software. With this solution there is no middleware necessary. The consequent use of XML allows communicating for the registration process via REACh-IT and the IUCLID 5 database. The solution provides advantages concerning the high scalability, platform and system independence, simple data exchange, partition of the application logic user interface and the complete integration of Java. In 2009 it is foreseen developing Web Services, enabling IUCLID 5 to directly communicate with other software applications.

The European Commission and other organisations, especially the OECD has triggered the activities, the methodology and the technology of the IUCLID 5 – application to use for other legislations of chemicals. Especially for pesticide and biocide the IUCLID 5 system could be a solution for efficient data collection and exchange of information between the parties. The IUCLID User Group Expert Panel under the steering of the OECD has to prioritise issues to be fixed and new functionalities to be implemented in the future releases of IUCLID 5.

REFERENCES

Boberski, C.: REACh – Handbuch, Leitfäden, Checklisten und Dokumente für die praktische Umsetzung der neuen Chemikalienverordnung. Forum-Verlag Herkert GmbH, ISBN 978-3-86586-126-9, München, 2008.

ECHA: Präsentationen zum REACH-IT Stakeholder Workshop, Brüssel, 2008. http://echa.europa.eu/news/20080215workshop_en.asp

ECHA: Präsentationen zum REACH-IT Stakeholder Workshop. Helsinki, 2008. http://echa.europa.eu/news/20080507workshop_en.asp

ECHA: REACh-IT – WebSide http://echa.europa.eu/reachit_en.asp, Helsinki, 2009.

Knetsch, G.: REACh – Workflows and Software Tools for the Process of Registration, Evaluation, Authorisation and Restriction of European Chemicals. In: Proceeding of the 22th International Conference on Informatics for Environmental Protection (EnviroInfo 2008), Shaker Verlag, Lüneburg, 2008.

TechniData: Präsentationen zum REACH-IT Stakeholder Workshop. Ispra, 2006.

Umweltbundesamt: REACh in der Praxis, http://www.reach-konferenz.de/, Berlin, 2008.

Umweltbundesamt: Jahresbericht des Umweltbundesamtes 2006, http://www.umweltdaten.de/publikationen/fpdf-l/3307.pdf, Berlin, 2007.

Umweltbundesamt: Informationsportal zu REACh http://www.reach-info.de/index.htm, Berlin, 2008.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pilmmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Conversion of Czech National Emission Inventory for Input to Air Quality Models

J. Liczki, J. Resler, P. Krč and K. Eben[a]

[a] Institute of Computer Science, Academy of Sciences CR, Prague, Czech Republic
(liczki/resler/krc/eben@cs.cas.cz)

Abstract: The Czech national emission and air pollution sources inventory (REZZO) is a large database covering anthropogenic emissions produced in the Czech Republic. It is the most natural and legitimate source to be input into advanced air quality models (AQMs), e.g. CAMx, CMAQ. The step which converts inventory data into emission inputs needed for AQM is usually called "emission model". Sparse Matrix Operator Kernel Emissions (SMOKE) system was developed as a powerful tool that enables development of emission models. It has been used for converting US inventory emissions into AQM format. This preprocessor is however restricted to US standards. In this contribution we show that after some adaptations, SMOKE can be used for building emission inputs for the Czech Republic based on the REZZO inventory. The issues encountered in processing emission inventories and in the setup of some building stones of the emission model are discussed. Converted REZZO emissions for point and area sources for a grid with 3km horizontal resolution are presented.

Keywords: Emission Inventories; SMOKE; Air quality modelling.

1. INTRODUCTION

Air quality modelling plays a significant role in enviromental sciences. Recently, advanced Air Quality Models (AQMs) have been developed; they are used for many purposes. Emissions data (both anthropogenic and biogenic) are one of the most important inputs for these AQMs. European anthropogenic emission inventory data are gathered and processed from national inventories into a central database (EMEP). These data can be used for AQMs over the territory of Europe. The EMEP grid resolution is coarse (50 km x 50 km) and an important portion of information is lost. A way to avoid these inaccuracies on the input to AQMs is the use of national emissions inventories.

Emission inventories are typically available with an annual-total emission values for each emission source. The AQMs, however, require emission data on an hourly basis, for each model grid cell (and perhaps model layer) and for each model species. It is necessary to transform emission inventories before using them as inputs to AQMs. The Sparse Matrix Operating Kernel Emissions (SMOKE) modeling system is a computationally efficient and flexible tool for converting an emission inventory data into the formatted emission input files required by an AQM, e.g. CAMx, CMAQ.Since

SMOKE was created at University of North Carolina at Chapel Hill it uses US standards and therefore it is not straight applicable as an emission processor for European/national emission data. In the following, after quick overview of SMOKE and REZZO, we disccus issues encountered during preparation of the inventories from REZZO for an input to SMOKE.

2. SMOKE - OVERVIEW

SMOKE supports area, mobile, and point source emissions processing and also includes biogenic emissions modeling through both BEIS2 and BEIS3 system. Biogenic emissions are transformed in a different way than the anthropogenic ones and we will not deal with them henceforth, they are a matter of a future work.

In general, a run of SMOKE for point, area and mobile sources covers five main steps:

Input of raw inventory data. Annual totals emissions invetories are on input into SMOKE. Individual sources are defined by their *characteristics* (e.g. state/county FIPS code, Source Clasiffication Code (SCC)) that ensure the uniqueness and *attributes* (e.g. flue gas temperature for point sources) that supply an additional information.

Temporal processing. The temporal processing operation applies factors based on the source characteristics to the emission data from the SMOKE inventory files. These factors can include monthly, weekly and diurnal temporal profiles. The resulting emissions data vectors contain hourly emissions for the inventory pollutants.

Splitting inventory pollutants into chemical species. Based on inventory pollutants on input, during the chemical speciation step SMOKE uses the instructions for supporting a specific chemical mechanism set by user and generates model-ready emissions for that specific chemical mechanism.

Spatial processing operation (gridding). It combines the grid specification for the AQM domain with source locations from the SMOKE inventory file. The gridding step is different depending on the type of source being processed

The merge step combines the hourly emissions, speciation matrix, and gridding matrix to create model-ready emissions.

All these steps can be considered as matrix - vector manipulations, moreover the matrices are sparse. SMOKE was designed to take advantage of these facts in order to achieve high efficiency on modern computer architectures.

3. INPUT DATA

The REZZO inventory is administered by the Czech Hydrometeorological Institute. It contains annual-total anthropogenic emissions for the Czech Republic, and it is divided into four categories according to the type of emission source:

REZZO 1 - large sources of emissions from combustion processes and technologies. Besides annual-total emission, additional stack parameters such as longitude, latitude, stack height and diameter, stack gas exit temperature are provided.

REZZO 2 - mid-sized sources of emissions from combustion processes and technologies. Only annual-total emissions and a less precise location are available. Typically, district heating plants are included.

REZZO 3 - small sources of emissions from combustion processes and technologies. In particular, annual-total emissions from local heating for catastral units are given here

REZZO 4 - transportation. Annual-total emissions from onroad and nonroad mobile sources for catastral units.

Below, we describe the issues encountered when preparing SMOKE-ready files using REZZO inventories together with associated data.

4. PREPARATION ISSUES

4.1 Preliminaries

Three different categories of anthropogenic sources are recognized by the SMOKE processor: point sources, area sources and mobile sources. Since REZZO 1,2 can be considered as point sources, REZZO 3 as area sources and REZZO 4 as mobile sources it seems to be natural to apply SMOKE as an emission inventory processor for REZZO. To do that we were faced with the problem of accomplishing the following tasks:

First, it was necessary to convert Czech territorial structuring into the format consistent with that one in the US. Also, US Source Classification Codes (SCC) classifying different anthropogenic emission activities were replaced by European Selected Nomenclature for Air Pollution (SNAP).

Second, auxiliary files were generated as a first step towards building a realistic emission model. Temporal (monthly/weekly/diurnal) allocation profile files are based mainly on information about production cycles for particular SNAP categories, chemical speciation files are defined by the use of chemical mechanism chosen. We have prepared speciation files for both CB-IV and SAPRC 99 mechanisms.

When these preliminaries were carried out, emissions inventories from REZZO could be transformed.

4.2 Point sources

REZZO 1 and REZZO 2 contain typical point sources, emissions are apportioned to the grid cell intersecting the point on the basis of knowledge of latitude/longitude position of individual sources. A special step is needed that depends on the type of AQM. In particular, when inputs for CMAQ are generated, meteorological data are needed. These data enter a plume-rise model used for distributing emissions into appropriate vertical levels.

4.3 Area sources

Emissions from REZZO 3 were considered as area sources. For area sources, catastral unit-total emissions are spread among the cells intersecting the catastral unit through the use of gridding surrogates (note that the surrogate is made for a specific grid). The surrogates are created using the information that is available at a finer resolution than the catastral unit data. Typically, this information comes from census returns (in our case supplied by the Czech Statistical Institute). Since REZZO 3 contains emissions from local heating, the set of data on housing from census returns was applied to build area surrogates. The data set describing all buildings in the Czech Republic (about 2.5 millions) includes in addition to the exact longitude/latitude position the information about the number of flats, the way of heating (local or district), the type of fuel (gas, electricity, coal, etc.), type of building (factory, school, hospital, summer cottage, block of flats, etc.). Based on these pieces of information, only buildings with residential heating were used for construction of the surrogates (since the emissions from heating the other buildings are covered by point sources in either REZZO 1 or REZZO 2). The factors that characterize emissions distribution among the grid cells were calcutated according to the technique presented below.

- The location in the grid is assigned to every building using both the grid definition and the lat/lon building position.
- Let us suppose that we have catastral unit k and grid cell (i,j) Then $emis\_fact(k,i,j) = \sum_{l \in L \bigcap P} num\_flat(l) / \sum_{p \in P} num\_flat(p)$, where P={set of buildings in catastral unit k}, L={set of buildings in cell (i,j)} and $num\_flat(l)$ represents the number of flats in the building l.
- The amount of emissions in the cell (i,j) is calculated as $\sum_{k \in K} = emis\_fact(k,i,j) * emis(k)$, where K={set of catastral units that intersect cell (i,j) } and $emis(k)$ is the k-th catastral unit emission total

Figure 1 illustrates the effect of the surrogates application for area sources.

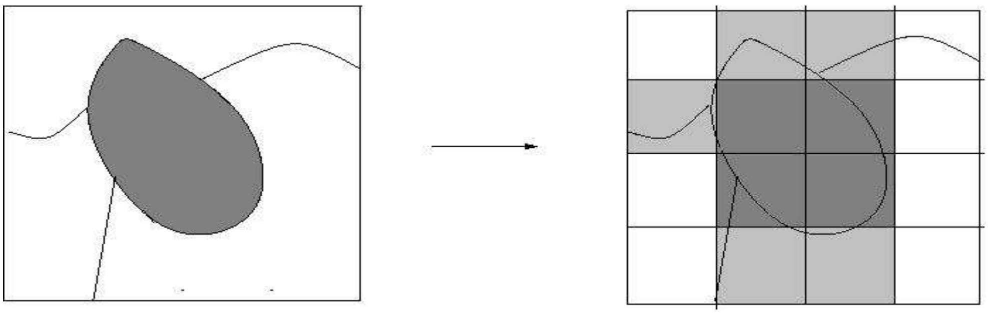


Figure 1: The grey colour represents the catastral unit-total emissions (left), various shades of grey represents various emission values on grid cells (right)

4.4 Mobile sources

There is a special SMOKE module MOBILE6 that allows an extremely precise processing of mobile emissions. Unfortunately, the transportation data required for this purpose are not available at the moment, although some preliminary steps have been made towards using this module. Therefore in the near future we shall use REZZO 4 emissions, which will be processed in a similar way as REZZO 3. The surrogates will be created from the information on the public road system.

5. RESULTS

Figure 2 and Figure 3 show typical output from SMOKE for area sources at 0 am and at 8 am, respectively. In Figure 5 and Figure 4 typical SMOKE output for the the largest point sources at 8 am is presented using different scales. The influence of fine resolution is apparent in both cases: point source emissions are located quite precisely, area source emissions are well distributed according to the surrogates.

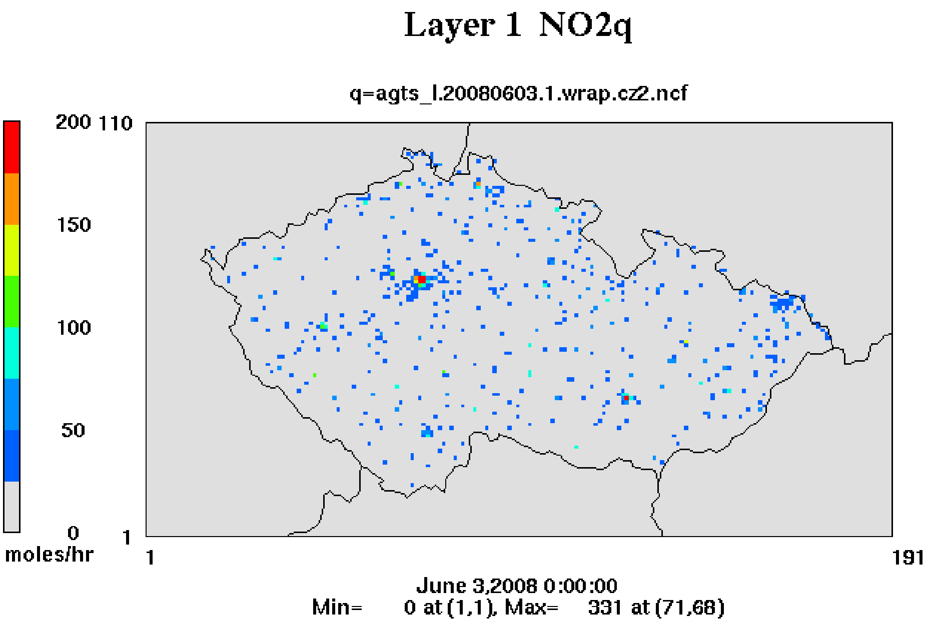


Figure 2: Example of gridded, speciated, temporal allocated area emissions (REZZO 3) when housing surrogates are used on the 3km resolution grid

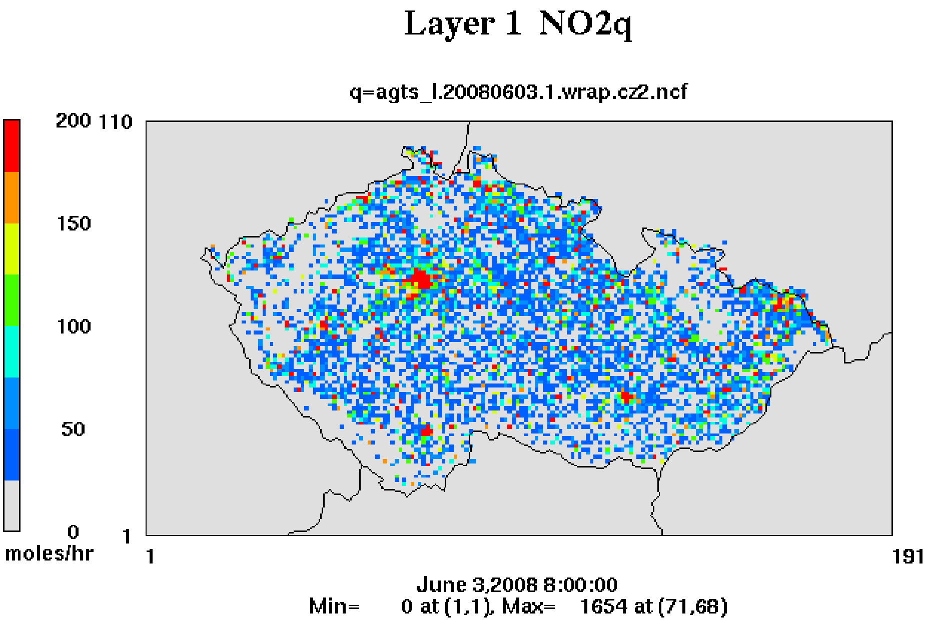


Figure 3: Example of gridded, speciated, temporal allocated area emissions (REZZO 3) when housing surrogates are used on the 3km resolution grid

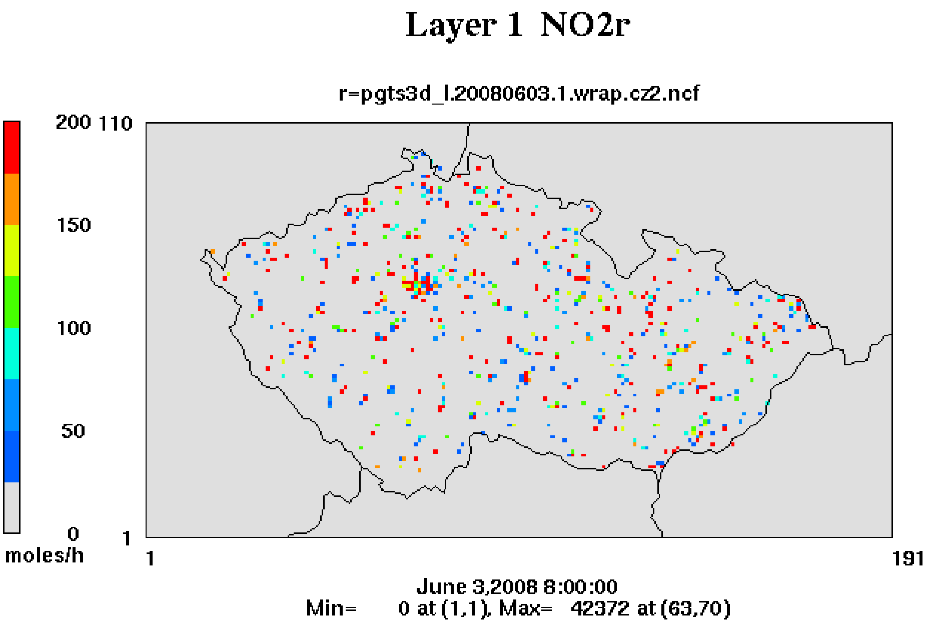


Figure 4: Example of gridded, speciated, temporal allocated point emissions (REZZO 1) on the 3km resolution grid

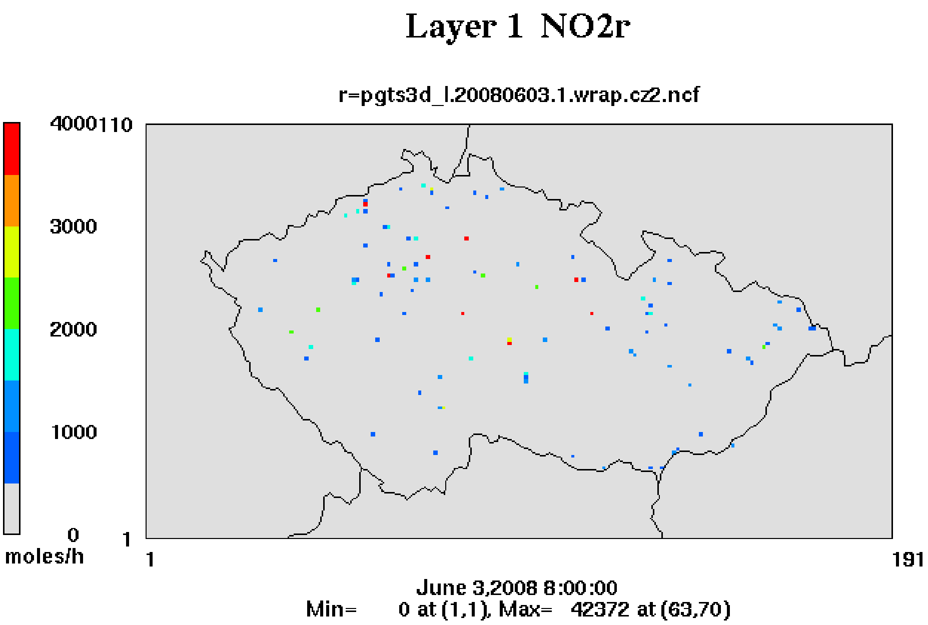


Figure 5: Example of gridded, speciated, temporal allocated point emissions (REZZO 1) on the 3km resolution grid

6. CONCLUSIONS

Emission inventories REZZO 1-3 have been processed with the SMOKE emission preprocessor, a simple emission model has been used and model-ready emissions for the AQ models CMAQ or CAMx have been produced in a sufficient horizontal resolution.

This improves considerably the precision of inputs for the AQ models. Moreover, flexibility of the SMOKE preprocessor and software tools developed so far give a good starting point for building and improvement of more complicated sofisticated models.

REFERENCES

Czech Hydrometeorological Institute, Czech National Register of Emissions (REZZO), 2004

Czech Statistical Institute, Census returns data set, 2001

SMOKE v2.5 User's Manual, University of North Carolina at Chapel Hill, 2008, *http://www.cmascenter.org*

U.S. Enviromental Protection Agency (EPA), Models-3 - Community Modeling and Analysis System, *http://www.cmascenter.org*

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Green IT Solutions

Georg Meixner
IBM CEEMEA, (meixner@at.ibm.com)

Abstract: The challenges have become clear: the need for clean water and air; affordable and reliable delivery of energy; the dwindling supply of fossil fuels; the reality of climate change and its implications for future generations. At IBM, our approach is twofold: we are working to make our existing products and processes more efficient for both the environment and for business, while also developing new innovations that can accelerate the adoption of products and services that have lesser environmental impact. Today's energy- and climate-related issues are at the top of our strategic agenda. We recognize that information technology plays an extremely important role in helping solve the myriad of ecological challenges faced by the global society—such as conserving our scarce resources even as global demand skyrockets, reducing pollution, minimizing the environmental impact of our activities, and enabling safe and renewable alternative sources of energy.

Keywords*:* Energy efficiency; Clear air; Climate-related issues; Reduce costs

1. HIGHLIGHTS

- Reduce overall power and cooling costs in your data centre or server room without affecting system performance, efficiency and sustainability
- Improve server and storage usage while minimising the physical space required by IT assets
- Gain compliance with emerging industry and government regulations while maintaining a flexible, resilient IT infrastructure.

When organisations go ‘green,’ the immediate perception is that they have implemented environmentally friendly practices. However, the rewards of minimising power consumption go beyond preserving the environment to creating more efficient systems and hardware, improving processes and reducing costs. They are also frequently surprisingly easy to achieve and maintain.

- Rising overall energy costs and demand
- Concern over outages and network interruptions due to power surges or spikes affecting business
- Wasted network resources and server space not being fully utilised
- Government-imposed regulations and standards becoming more strict
- The need for an environmentally safe way to dispose of old software and hardware.

It is not unusual for environmental issues to spark ‘hot’ debate in the boardroom these days, but for many businesses, what was once just talk is rapidly turning into reality. Organisations large and small are becoming increasingly concerned over the rapid growth of energy costs. On November 7, 2007, IBM released the results of a global survey of nearly 1,400 small and mid-sized businesses (SMBs) identifying energy costs as the ‘biggest cost increase’ for them over the past two years, higher than even healthcare. At the same time, rates of consumption are spiraling upward, driven in part by the expanding role that IT plays in business. In fact, energy usage has escalated from just another line expense to a core corporate issue.

Furthermore, as federal and local governments intensify focus on this area, an abundance of new regulations could adversely affect business operations, making the risk of non-compliance an added threat.

2. WHAT DOES 'GREENING YOUR IT' MEAN FOR YOUR BUSINESS?

Being green extends beyond hardware assets to include facilities design and operations. For a mid-sized business with a large computer room or full data centre, 'going green' means more than just turning off the lights or recycling cans. It is a fundamental shift in your approach to your IT infrastructure. Becoming energy efficient starts with a comprehensive assessment of your IT environment. Familiarity with conditions in your physical plant – both the power consumed and heat generated by your IT assets and your heating and cooling systems – is critical to your ability to achieve the right balance.

Today, innovative IT solutions can make it easier for your business to become energy efficient by helping you:

- Lower energy costs while sustaining IT performance
- Monitor power surges and shortages that can affect overall business
- Minimise risk of overload in the data centre by maximising server utilisation
- IBM technology can help reduce overall power consumption
- Virtualization can reduce the number of servers needed, while maximising network resources
- IBM Rear Door Heat eXchanger can help ease the burden on existing air conditioning units and can potentially lower your energy costs by up to 15 percent
- IBM services can help optimise your IT solution investment by providing asset disposition services and certified used equipment.

3. AN APPROACH THAT WORKS

Costs and environmental concerns are forcing businesses to rethink the way their core business operates. The need for access to the right information, applications, processes and people hasn't changed, but new ways of supporting this infrastructure are needed as energy use and power costs become critical issues in your day-to-day operations.

IBM's green IT efforts aim to remove the barriers that often prevent SMBs from adopting energy saving, environmentally friendly IT practices by making them financially attractive and easy to implement. IBM encourages SMBs like yours to take the following steps toward improving the energy efficiency of IT systems:

- Make energy efficiency a priority when purchasing new IT systems
- Consolidate with virtualization technology. You can achieve a smaller IT footprint by pooling resources from multiple systems so that you get more computing from fewer systems
- Have a plan for how to dispose of IT equipment in a secure and sustainable manner. Certified recycling services will ensure that data is wiped clean and the systems are reused, resold or recycled, keeping harmful materials out of landfills.

Energy Efficiency solutions include services, systems and software that can help power your green initiatives for your data centre or server room and ensure that you're getting maximum efficiency from the smallest amount of energy possible. These range from innovative power and cooling technologies to eco-friendly asset disposal services.

When used as the foundation for an efficient infrastructure or to enhance existing supported environments, IBM's five-step approach to assessing the energy efficiency of your current

IT infrastructure can provide your business with new and innovative ways to meet overall energy management goals. The steps in this approach are:

1. Diagnose
2. Build
3. Virtualize
4. Cool
5. Manage and maintain.

4. FINANCE

When you choose IBM Global Financing as your commercial financing partner, you will be in a stronger position to meet the needs of your business. The combination of extensive resources, global reach and industry-specific expertise makes IBM a great partner for all your commercial financing needs

5. GROW YOUR BUSINESS IN A GREEN WAY

With Energy Efficiency solutions, you'll have the tools and awareness to understand the problems that your company faces when dealing with the issues of increased energy usage and environmental costs. Your business can continue to grow and innovate while your IT department optimises costs with solutions that are easier to implement and can have a dramatic effect on your overall business outlook.

IBM is dedicated to supporting businesses with their green initiatives, helping them to reduce energy consumption, achieve power and cooling capacity goals and regain control over their infrastructure. For example:

- IBM has been an innovator of energy-smart technology for over 40 years – from breakthroughs in mainframe cooling to the development of the powerful Blue Gene computer that delivers the most performance per kilowatt of power consumed
- IBM 'practices what it preaches.' Between 1990 and 2005, IBM's global energy conservation actions reduced or avoided CO_2 emissions by an amount equal to 40 percent of its 1990 emissions
- IBM has always been an innovation company that focuses on creating new technology and business capabilities that tackle important problems.

IBM technologies such as Rear Door Heat eXchanger, automated system standby, advanced server virtualization and high-density computing offer a complete green IT package that is complemented by the Global Asset Recovery Services program.

REFERENCES

IBM - Energy, the environment and IBM, *http://ibm.com/green*

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Environmental Economic and Natural Disasters

E. Pajorova[a] **and L. Hluchý**[a]

[a]*Institute of Informatics, Slovac Academy of Sciences Dúbravska cesta 9, 845 Slovac republic (utrrepaj@savba.sk)*

Abstract: Every year, forest fires, floods and landslides cause enormous damage of vegetation and fauna, environment and property and bind significant human resources. Particularly in national parks and natural reservations, unique areas of high degree of protection can be devastated by fire. For instance, during the destructive forest fire in the Slovak Paradise National Park (Slovakia) in 1976, very unique vegetation was destroyed in the Kyseľ Gorge, where the recovery into the former state will last 200 years [Juhas 2000]. Till now (thirty three years after the fire), this locality is closed for tourists because of the vast damages. Topic of a lot of projects is how to prevent such disasters. Our research in institute is oriented on GRID computing. A lot of international projects oriented on natural disasters utilize the grid computing and within grid solution raises requirement of visualization service for presentation of the intermediate or final results. Our basic aim of our research resolved in projects is the creation Visual service for the Modelling and 3D Rendering of Natural Disasters, before Fire, floods and landslides. Changing input data of Fires spread using to generate new outputs very quickly. Grid computing on a lot of Clusters and On-line 3Dvisualisation service can allow new scenes of Natural disasters spread. Outputs are using for far adjustment to liquidation the Natural Disasters.

Keywords: Environmental Economic; Natural Disasters; simulation; modelling and visualization; Grid computing.

1. INTRODUCTION

3D_visualization service for animation Natural Disasters should integrate visualization requests of any kind of application solved in our institute and before solved in international projects oriented on environmental problems. The natural disasters like fires and floods or landslides become subject of science in research institutions more and more frequently. Many applications from this area are using different kinds of simulation tools, which are producing output data for displaying the results of the computation. The purpose of the visualization service is to model and display results of various simulations of natural disasters like are fire spread in time, fire intensity, flood velocity, landslide activity etc. Such service requires unified standards like integration of input data formats and especially creation of unified visualization tool.

This paper is describing such a visualization service developed and tested in our institute, create a 3D visualization service for GRID oriented natural disasters applications. The purpose of 3D viz. service is to model and display intermediate or final results of various simulations of natural disasters like fire spread in time, its intensity and erosion or floods in time or landslides as well. The output of the service is various scenes of terrain by different simulation outputs. Output of the service can also be the files representing the virtual reality of natural disaster and also files, which are generated as input for VR-Systems. 3D service was tested with outputs from applications, which were solved in our institute and also by data from applications of the MEDIGRID project [MEDIGRID EU 6Rp].

Visualizations tool was tested on fire simulation in Marseille industrial port in the southeast. The second simulation was a large fire in the Krompla region [Glasa, 2006], which is part of the Slovensky Raj. Flood demonstrations are from Povazie, the region around the river Vah. The root of visual service is to create a modelling tool. A modelling tool consists of modules, which were necessary to create. Each module is a UNIX shell script in which is prepared the start of the executables. The modules are divided into three groups according to what kind of output 3D models the group is generating.

2. MODELS GENERATING

- Models of terrain or model of environment (Figure 1).
- Models of simulations (Figure 2). .
- Virtual reality models of terrain and simulation (fire spread, flood, landslide etc.) (Figure 3).

Figure 1. Terrain around the river Vah

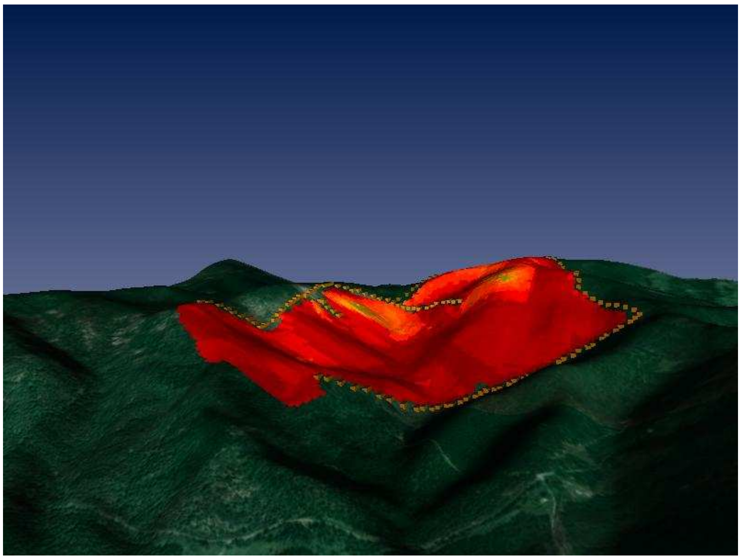


Figure 2. Fire spread simulation on Krompla region

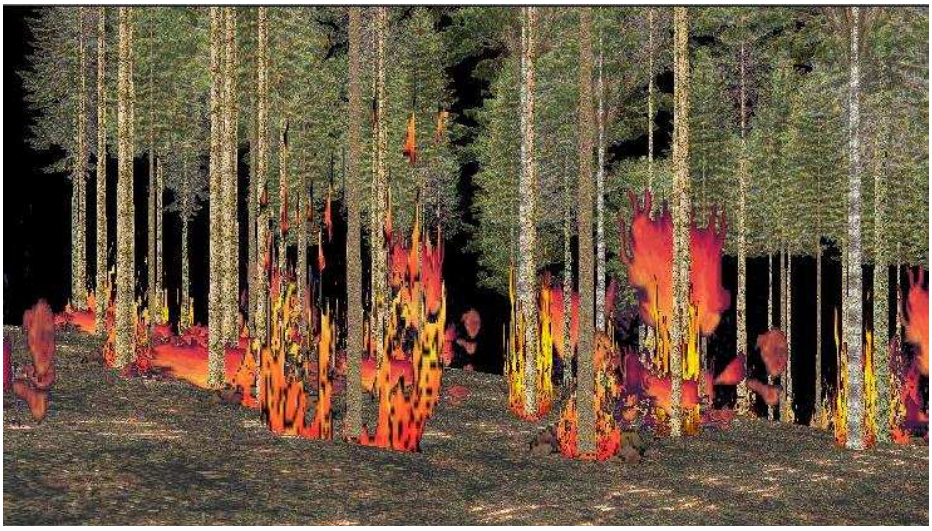


Figure 3. Virtual reality model - Fire spread

Natural disasters simulation is a very complicated, challenging problem sensitive to the input data required [Anderson, Ricards]. Therefore, intense research and development of sophisticated software systems and tools is extremely important for Natural disasters fighting management purposes. For example for Slovak forests, original methodology for forest vegetation classification and new fuel models have been developed and proper forest fire simulations related to the locality Krompla (National Park Slovak Paradise), where the large destructive fire appeared in 2000 and its reconstruction have been analyzed [Tucek 2004]. These efforts induced the need of better auxiliary tools for 3D visualization of obtained simulation results and for animation of the forest fire spread. In this paper, new 3D visualization technique for real forest fire simulation and fire behavior and for flood and landslide modeling is described. The importance is increasingly expanded for environmental problems [Viegas 2004].

3. ON-LINE VISUALIZATION AS A SEQUENCE APPLICATION IN GRID ENVIRONMENT

Grid-based applications that generate visualization outputs on-the-fly require a unified application framework. Our aim in this paper is to design a framework that will define a unified way of cooperation between Grid visualization applications and visualization clients. Computations in the Grid environment being exceedingly long, the user naturally wants to see the intermediate results and requires means to modify the running computations.
With regards to the Grid computational environment, it is advantageous to divide the performation of Grid applications into three separate phases.

1. Fetching the data from a simulation application
2. Transforming the data into a displayable form and the creation of virtual scenes
3. Displaying the scenes

3.1 Fetching the Data from Simulations

The first, computational phase is performed in the Grid. The whole application consists of a large number of consecutive series of jobs. The output from a given series becomes the input for the next series. To perform it, in particular we used the EGEE II, its virtual organization VOCE for central Europe, and Trigrid for south Europe. The Storage Element is used as a data server, from which the computational agents *(workers)* download data for processing. The worker processes data in a loop until its time quantum is over, or until all data are processed. The control script *(job controller)* runs in the background on the User configuration. The user must prepare the data on the Storage Element and provide the configuration file before launching the job controller. Data can be added even while the application is running (providing the relevant rules are kept). During its run the user can observe the contents of directories with input and output data on the Storage Element. The simulation outputs are treated as input data by the visualization process. The job controller starts their processing by launching the visualization tool. This is the subject of Phase 2.

3.2 Transforming the Data into a Displayable Form

The aim of this phase is the transformation of the data into a displayable form and the creation of virtual scenes according to the requirements of the client. This phase is carried out by a visualization tool, which is together with its controlling agents, is the crux of this article. Apparently, the best approach is to create a separate visualization tool for each type of simulation. Only closely related applications, for instance different simulations of environmental phenomena, enable an usage of a shared control script.The user is free to decide in which phase of the computation the visualization tool is activated. Alternatively, the tool can be run after the whole computation is completed. While the computations are performed, the user can observe the contents of directories with input and output data on the Storage Element. A simulation output is in a certain phase considered to be the input for the visualization process. In the next cycle, the user prepares a new configuration file and runs the visualization tool as a new application.

3.3 Displaying the scenes

The purpose of this final phase is the display and analysis of output files from the second phase.This phase takes place on the client graphical device or in Grid to instal plugin.
In new design for On line visualisation tool is included also online approach for rendering

4 NATURAL DISASTERS APPLICATIONS SOLVED AS A PARAMETRIC STUDY

Lot of applications of Natural Disasters are solved as a typical Parametric studies. Parametric study execute one application many times with different sets of input parameters. It is considerable for computing Natural Disasters Grid applications actually and very quickly computing with changes parameters, which can be for example weather, wind, fuel etc. To obtain outputs in least time, we designed tool for on-line visualization in which started visualization as a sequence application.
The process of On-line visualization described schema see (Figure 4).

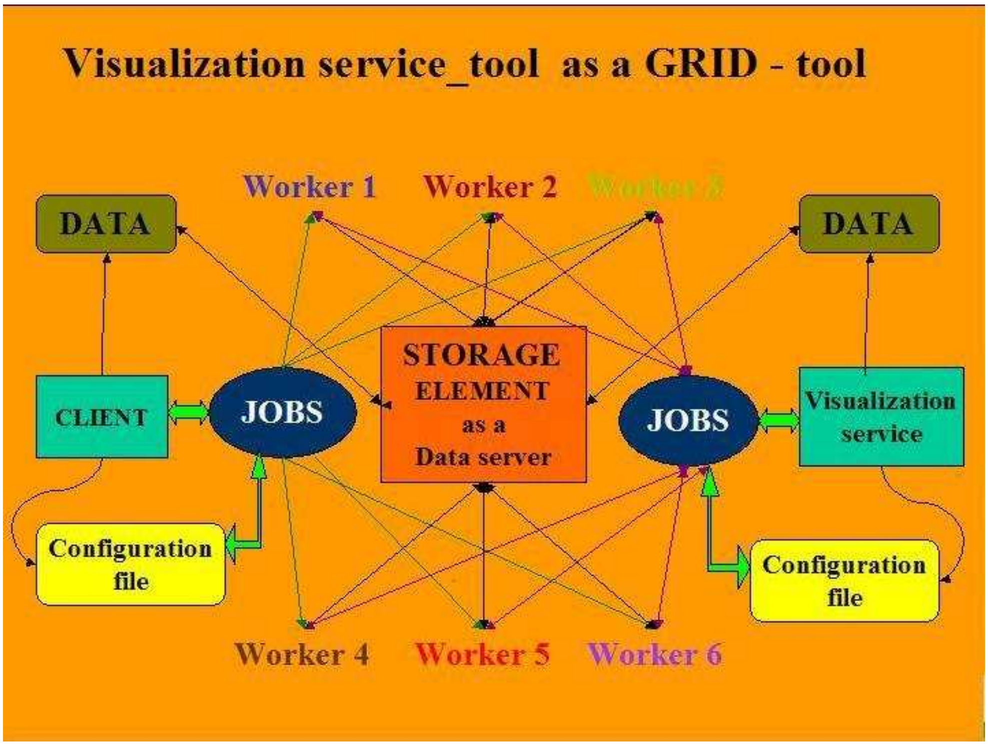


Figure 4. Visualization service tool as a GRID tool

Grid-based applications that generate visualization outputs use a unified application framework, which can to guard the unified way of cooperation between Grid visualization applications and visualization clients. User naturally wants to see the intermediate results and requires means to modify the running computations. All client requirements are respect in our design of On-line visualization, which is activated as a sequence Grid application. The process is control in some consecutive steps.

- Submission: a bulk of the linked jobs was emplaced on the Storage element. SE as a data server
- Script started on the User Interface controlled another script on Computing Element
- Workers worked till inputs stored on the SE was completedClient ask for visualization is as a Visualization client
- Outputs data on the Storage element are as a inputs data for visualization jobs
- Configuration file will be included by Visualization tool for the application
- Can be activated control script
- Workers modified data to the visualize formats

5 CONCLUSIONS

Our aim for the future is to propose such a Framework, which included sophisticated and generic tools for computing, before Grid solution of computing, and the On-line visualization tools, which can be activated as ***a sequence application in Grid environment***. Framework will be using for far adjustment to liquidation Fires or floods and Landslides more and more. We can allow them for the solution of applications, which are deal with European and global world wide problems with Natural Disasters.

ACKNOWLEDGEMENTS

This work was partially supported by Science and Technology Assistance Agency under the contract No. APVT-51-037902 and by the EU project MEDIGRID EU 6FP RTD GOCE-CT-2003-004044. And also actually is supported by VEGA project No. 2/0211/09.

REFERENCES

Juhas, F.: Forest Fire Slovak Paradise National Park (in Slovak), Report and Documentation of Department of Forest Fire Services, Spisska Nova Ves (2000).

Viegas, D.X., Bibeiro, L.M., Silva, A.J. and Palheiro, P.: Analysis of S. Domingos accident, Proc. of the 4th Int. Conf. on Forest Fire Res. and Wildland Fire Safety, Luso, Portugal (2002) 18–23.

Anderson, H.E.: Aids to determining fuel models for estimating fire behavior. USDA For. Serv. Gen. Tech. Rep. INT-122 (1982).

Glasa, J., Halada, L.: Application of envelope theory for 2D fire front evolution. Int. Conf. on Forest Fire Research, Coimbra (2006).

Richards, G.D.: An elliptical growth model of forest fire fronts and its numerical solution, Int. Journal for Num. Methods in Eng., 30 (1990), 1163–1179.

Tuček, J., Schmidt, M., Celer, S.: Klasifikácia vegetačného krytu vo vysokohorských podmienkách z materialov DPZ s vysokým rozlíšením pri uplatnení apriorných poznatkov. Acta Facultatis Forestalis, vol. XLVII (2004), 89–102.

Viegas, D.X.: Surrounded by fire (in Portugal), Editorial Minerva, Coimbra (2004), 283 p.

MEDIGRID EU 6FP RTD GOCE-CT-2003-004044. Hluchý, L., Habala, O., Nguyen, G., Šimo, B., Tran, V., Babík, M.: Grid Computing and Knowledge Management in EU RTD Projects of IISAS.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Assimilation of satellite and ground level data into air quality models

J. Resler, K. Eben, P. Juruš and P. Krč

Institute of Computer Science

Academy of Sciences of the Czech Republic, Prague, Czech Republic

Abstract: Tropospheric columns of various species retrieved from satellite instruments bring a new quality into data assimilation systems for air quality models. In our experiments we assimilate tropospheric columns of NO_2 retrieved from the instruments GOME2 and OMI together with ground-level observations of NO_2. The 4DVar method has been used to optimize both initial conditions and suitable parameters of the emission model. A forecast experiment has been performed and the contribution of data assimilation to the one-day-ahead forecast has been studied.
Keywords: data assimilation, air quality models, satellite data, GOME2, OMI, IASI

1. INTRODUCTION

During the last decade a number of satellite instruments started to provide information on atmospheric chemical compounds. In particular, tropospheric columns of nitrogen and carbon oxides and ozone, aerosol optical depth and other products emerged. As a straightforward application, global trends in atmospheric pollution and greenhouse gases concentration have been investigated, based on yearly and monthly averaged satellite data. Using satellite products for short term prediction, inversion modelling of emissions and for simulation and modelling in finer horizontal resolution is less straightforward. Any data assimilating system attempting at ingestion of satellite data has to cope, among others, with the following specificities:

- instruments like OMI, SCIAMACHY, GOME2, IASI or MODIS are on board of polar orbit satellites, which means that they pass over a distinct region roughly at the same time each day. In order to obtain daily profile of concentration, one has to combine the information from different satellites and from other sources of information.
- satellite pixels may be large in comparison with model resolution.
- the data produced by the satellite instruments usually have the form of a vertical column or profile of a species, retrieved from the measured spectra. Thus the observation operator, which connects the model state variables to the observations, has the form of an integral which depends of some additional properties of the atmosphere. The retrieval process itself is a complicated procedure. All these facts contribute to measurement errors which have to be taken into account in data assimilation routines.
- tropospheric column is derived from total column by subtracting the stratospheric part. This fact together with the lower sensitivity of satellite instruments in the troposphere introduces further errors, sometimes very large.
- cloud presence causes other errors or makes the tropospheric column totally unavailable.

Despite of these difficulties, satellite observations bring a new quality to our sources of information. They can help to correct substantial model bias in higher atmosphere and they bring high spatial coverage and measurements over oceans and remote areas. In this sense, the satellite observations can be considered as a complement to the in-situ and aircraft observations.

We used for our simulations the numeric weather prediction model WRF and chemical transport models CMAQ and CAMx. Both model pairs have been configured for three nested domains with horizontal resolutions 27, 9 and 3 km.

2. ASSIMILATION OF SATELLITE COLUMNS AND IN-SITU OBSERVATIONS

In earlier work [Eben et al. 2005] it was found that correction of model concentrations achieved by assimilation of in situ observations is effective if we want to improve e.g. an exposure index estimated from long-term off-line simulations. For the purpose of prediction this kind of data assimilation has a limited benefit as long as model bias and errors in emission inputs are present. Another cause of this behavior is the lack of information for higher levels of troposphere.

In order to improve the forecast performance of the model, further sources of information are required and emission constraining appears to be necessary.

The 4DVar method is one of the most useful techniques of data assimilation. It seeks an optimal state vector by means of minimizing a cost function which penalizes the distance of the optimal state from observations and from the model values, giving both terms appropriate weights. The method can be naturally generalized so as to optimize emissions too, augmenting the state vector with emission parameters and adding a corresponding term to the cost function.

We used a similar approach as [Elbern et al., 2007]. As a first step, the variability in emissions has been roughly parameterized by allowing for an emission factor, specific for each gridpoint. This leads to the discrete formulation of the cost function

$$J(c_{0,}e) = \frac{1}{2}(c_0-c_B)^T B^{-1}(c_0-c_B) + \frac{1}{2}(e-e_B)^T K^{-1}(e-e_B) + \frac{1}{2}\sum_{i=1}^{N}(y_i-H(M(c_{0,}e,t_i)))^T R^{-1}(y_i-H(M(c_{0,}e,t_i))) \quad (1)$$

where

- $c_{0,} c_B$ are optimized, resp. first guess concentrations in time t_0
- e, e_B are optimized, resp. first guess emission multiplicative factors
- $c = M(c_{0,} e, t_i)$ are the modelled concentrations
- H is the observation operator
- y are the available observations, both satellite-retrieved columns and in situ observations
- B, K and R and the covariance matrices for initial conditions, emission factors and observations errors

The matrices B a K are taken diagonal (some tests were done with covariance matrices constructed by means of a diffusion operator). In the construction of the matrix R, estimates provided by the TEMIS service have been utilized for determining error variances of satellite columns.

The observation operator H has two separate versions, one for treating tropospheric columns obtained by the retrieval process and the other one for handling in-situ observations. The first one is computed by numerical integration and accounts for the spatial intersection of the satellite instrument pixel with grid cells and for the influence of the retrieval process given by the averaging kernel operator [Eskes and Boersma, 2003]. The second one is straightforward.

The term $y_i - H(M(c_{0,}e, t_i))$ is usually called innovation. It is useful to study innovations and their longer term averages to determine model bias and identify important locations in the domain. An illustration of such an analysis is in Fig. 1. for observations of tropospheric columns of NO_2 [Eben et al., 2007]. NO_2 has been retrieved by the instruments OMI and SCIAMACHY, its model counterparts are computed with the chemical transport model CAMx.

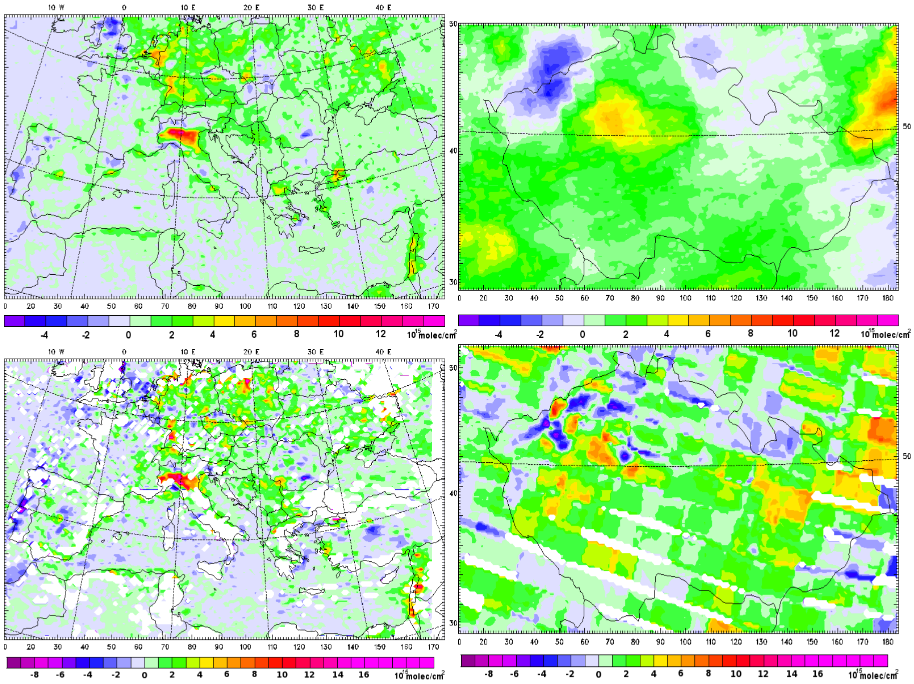


Figure 1. Mean innovation for the first (res. 27km, Europe) and third nested domain (res. 3km, Czech Republic). Upper part: OMI (overpass 13-16h), Lower part: SCIAMACHY (10–12h, much less data and coarser resolution). Prague region underestimated, North-west Bohemia overestimated by the model.

The gradient of J with respect to joint variable $(c_{0,} e)$ can be expressed as

$$\nabla_{(c_{0,} e)} J(c_{0,} e) = B_{-1}(c_0 - c_B) + K^{-1}(e - e_B) + \sum_{i=1}^{N} M^T H^T R^{-1}(y_i - H(M(c_{0,} e, t_i))) \qquad (2)$$

This gradient is calculated by means of the adjoint method. Therefore, a CTM capable of adjoint modelling had to be used. One of the few models with this property is the chemical transport model CMAQ. Its adjoint operator was being developed by the CMAQ community during recent years ([Hakami, 2007]). The implementation of the adjoint operator is done by similar means as in the model STEM (see [Sandu et al., 2003]). An experimental code has arisen in California Institute of Technology and Virginia Polytechnic Institute. This code contains the adjoint model for gas phase processes for the mechanism CB4. We have finalized the parallelization of the adjoint operator and contributed several technical improvements to the adjoint code so as to enable its use for real cases. Finally, observation operators for satellite columns and their adjoint have been implemented. We also have done some adaptations in order to constrain both initial conditions and emissions and we have built our 4DVar assimilation above this adjoint code.

A majority of our experiments has been done for CB4. Due to well know deficiencies of the CB4 chemical mechanism we have recently developed the adjoint for SAPRC99 chemical mechanism and performed first experiments with 4DVar.

The cost function J is minimized utilizing the L-BFGS-B algorithm [Zhu 1997].

3. THE ASSIMILATION EXPERIMENTS

In our experiments we used three nested domains (see Fig. 2). The outer domain encompasses most parts of Europe (horizontal res. 27km). It was used for obtaining realistic initial and boundary conditions for the assimilation run. The assimilation experiment was performed on a subwindow of the outer domain with 72x52 gridpoints. In some simulations, an intermediate domain (not shown in the figure, res. 9km) was used. It was also used for all WRF nested runs. The fine 3km resolution domain covers north-west part of the Czech Republic and adjacent regions of Germany (62x46 gridpoints). The emission data are based on the EMEP[1] inventory, the emission model was the same as in [Eben et al., 2005].

[1] *EMEP – European Monitoring and Evaluation Programme, www.emep.int*

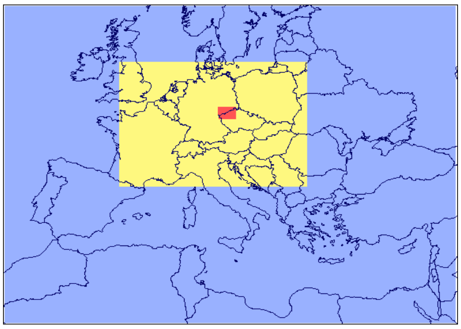


Fig. 2: Simulation domains: outer domain, assimilation domain, fine domain.

Tropospheric NO_2 columns are retrieved from measurements obtained by satellite instruments OMI and GOME2 (provided by the TEMIS service). The data contain all other necessary information, in particular the averaging kernel operator, the air mass factor etc. The necessary meteorological variables for constructing of the observation operator are taken from the WRF output files processed by MCIP3.

The overpass of the satellites above our domains occurs usually twice daily in the interval from 9am to 3pm. In Fig. 3 there is a typical example of the tropospheric column data (white squares represent cloudy conditions).

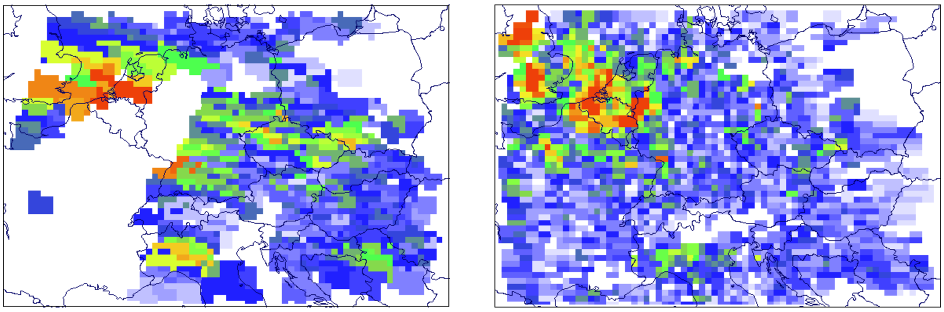


Fig. 3: An example of retrieved tropospheric column data from GOME2 (left) and OMI(right) for the coarse assimilation domain.

Altogether 280 background monitoring stations have been included into the experiment.[2] Their location in coarse and fine domain is in Fig. 4.

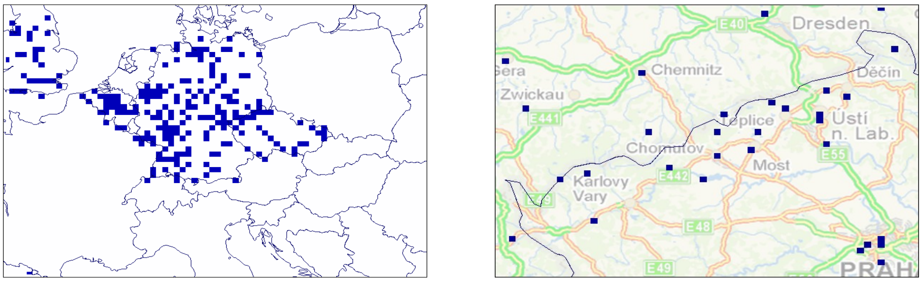


Fig. 4: Location of ground-level stations in the coarse domain (left) and in the fine domain (right)

We selected for our experiments an assimilation period of eight days from June 28 to July 5. A free run of CMAQ, 21 days on the outer domain, was performed to obtain reasonable initial and boundary conditions.

For each day sequentially, an assimilation run has been performed, assimilating both NO_2 columns and in-situ observations. A one-day-ahead forecast has also been computed, using emission factors and initial conditions obtained by 4DVar for the previous day.

Similarly as for CB4, an analogical experiment with the SAPRC99 mechanism has been performed. In this experiment observations of NO_2, NO and ozone have been assimilated.

4. RESULTS

During the first two days of assimilation, the estimated emission factors reached a fairly stable solution. After the two days the cost function showed only a decrease of a few percent, even though the number of observed satellite pixels (and thus the value of the cost function) varies due to changing cloud coverage. The map of emission factors for one

[2] *The selection of data has been dictated by their availability. We have used data from the UK and Belgium (supplied by courtesy of the European Environment Agency), Germany (provided by the Umweltbundesamt www.uba.de) and Czech republic (provided by the Czech Hydrometeorological Institute).*

particular day and the map of the corresponding gradient of the cost function with respect to the emission factors is shown in Fig. 5.

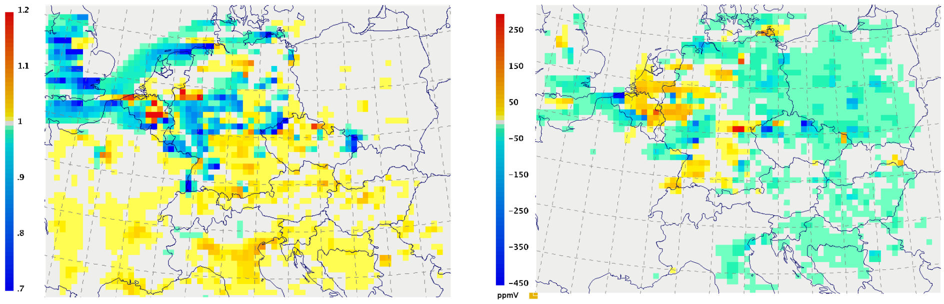


Fig. 5. Left: optimized emission factors for July 3, 2008 Right: Gradient of the cost function w.r.t. Emission factors, first iteration of the optimization process. July 3, 2008

The experiment with SAPRC99 and ozone data assimilation showed a similar behaviour. In Fig. 6 the emission factors for June 28, July1, July3 and July4 are shown. It is seen that both mechanisms give slightly different results. This is also due to additional information from ozone data in the SAPRC99 run.

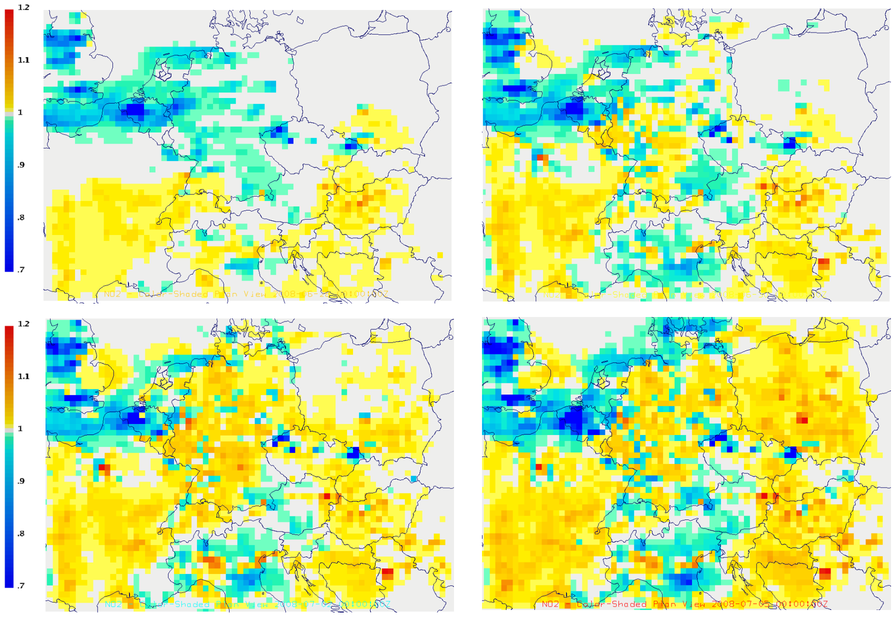


Fig. 6: Changes of optimized emission factors in the course of an assimilation experiment. Upper left: June 28, upper right: July 1, lower left: July 3, lower right: July 4, 2008

Fig. 7 depicts the changes between the free run and concentrations corrected by 4DVar, for a CB4 run and a SAPRC99 run. It is seen that both in situ observations and satellite columns contribute to the corrections. In particular, the corrections in France and Italy are induced by satellite observations only. There are, of course, significant differences between the two chemical mechanisms.

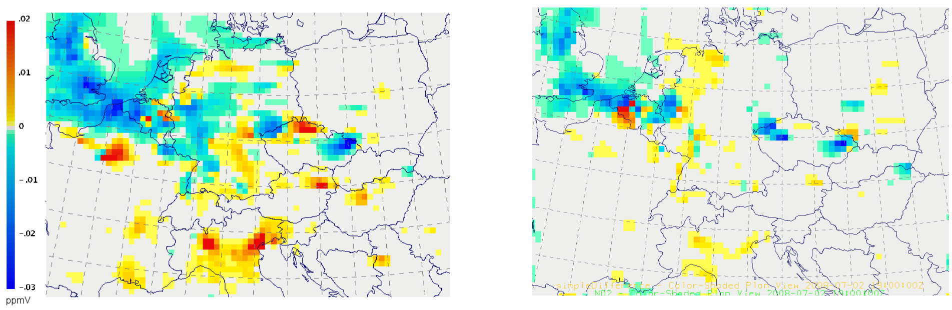


Fig. 7: Differences between optimized and referential concentrations of NO_2 , July 2, 2008 20:00. Left: CB4, right: SAPRC99

Table 1 contains results of the one-day ahead forecast for the CB4 run, compared with these of the free run. It is seen that the forecast from assimilated values outperforms the free run, but the bias of the forecast is larger.

We excluded small values of NO_2 from the evaluation, so that only values of NO_2 larger than 20, either in the observation or in the model, enter the evaluation.

No. of observations	Free run **mean residual**	Forecast **mean residual**	Free run **mean absolute res.**	Forecast **mean absolute res.**
58567	2.3	12.1	20.15	17.7

Tab 1. Mean residuals (differences of hourly observed value at a station and the model value) and mean absolute residuals for the free run and forecast from assimilated initial conditions and parameters. All values are in µg/m³.

In Fig. 8 we see one particular forecast for the fine domain, compared with the corresponding free run and the zoom of the coarse domain. Assimilation helps here to identify and localize sources in a better resolution (beyond the downscaling due to better orography and meteorology), even though the initial emission inventory is coarse.

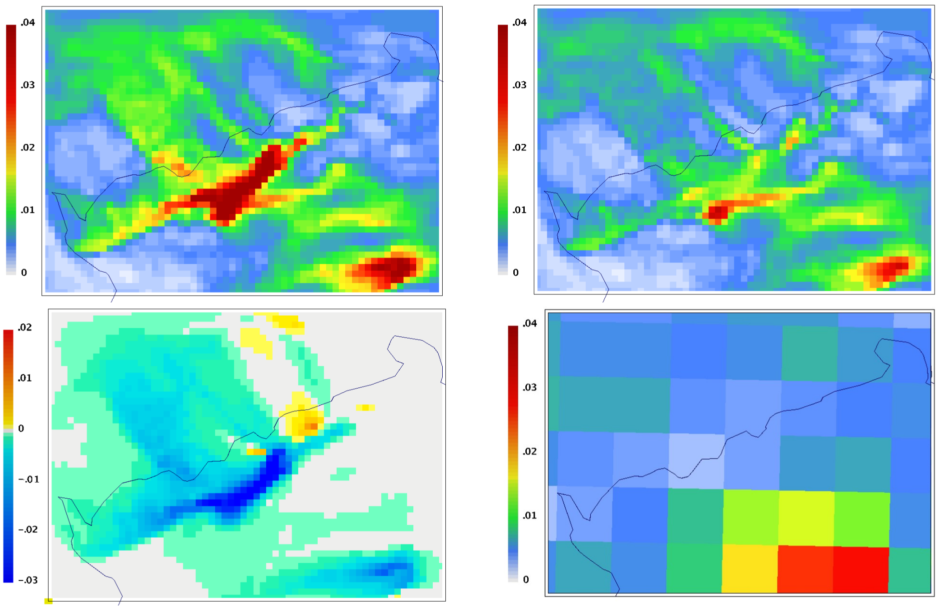


Fig. 8: Referential (upper left) and optimized (upper right) concentrations in ppmV for July 1, 2008 20:00 in the fine domain. The differences of the optimized and referential concentrations (lower left) and the optimized concentrations in the zoomed coarse domain (lower right).

In Fig. 9 there is an example of the time profile of concentrations for one of the stations in the polluted area, where the downward trend in emissions occurred during recent years. Although using an emission factor reduces the error substantially, the parameterization is too crude and a generalization is to be made, while keeping stability of the solution.

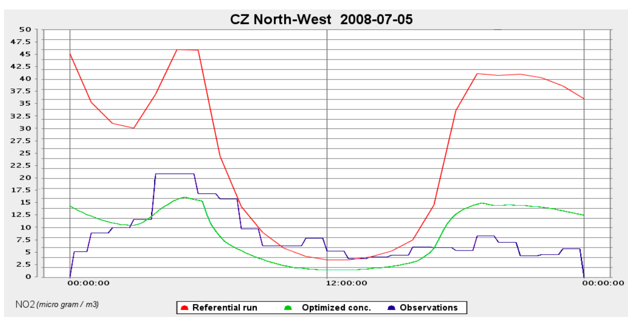


Fig. 9: Concentrations on obs. station Litoměřice (North-West Bohemia, the strongest changes by the assimilation).

5. CONCLUSIONS AND OUTLOOK

The assimilation experiment shows stability and good performance of the 4DVar method in our setting. A greater flexibility of parameterization of emission corrections will be needed since a simple multiplicative factor cannot achieve a better fit of time profiles. A subsequent statistical analysis of the relation between corrected emission fields and parameters of the emission model should lead to the adaptation of the emission model itself.

The improvement in forecasting NO_2 concentrations due to assimilation of in situ and satellite observations is moderate so far, but evident. A deeper investigation of complementarity between in-situ and satellite observations is required. This would result in a better description of vertical profiles of NO_2. The results of the test run with the SAPRC99 mechanism show high sensitivity of concentrations of some VOC's to changes in NO_2. These compounds have to be included into the assimilation process in the future. Finally, joint assimilation of other important species, in particular ozone, as well as more precise modeling of spatial covariances could improve the allover performance of the model.

ACKNOWLEDGEMENT

This work was supported by the Grant Agency of the Academy of Sciences of the Czech Rep. (grant No. 1ET400300414, framework „Information Society"), by the grant No. SP/1a4/107/07 of the Ministry of Environment of the Czech Rep. and by the Institutional research plan AV02 10300504 Computer Science for the Information Society "Models, Algorithms, Applications".

REFERENCES

Eben, K., Jurus, P., Resler, J., Belda, M., Pelikan, E., Krueger, B. C. and Keder, J. 2005. An ensemble Kalman filter for short term forecasting of tropospheric ozone concentrations, *Q. J. R. Meteorol. Soc. 131, No. 613*, 3313-3322.

Elbern, H., Strunk, A., Schmidt, H. and Talagrand, O. 2007. Emission rate and chemical state estimation by 4-dimensional variational inversion, *Atmos. Chem. Phys., 7*, 3749–3769.

Eskes, H. J., and Boersma, K. F. 2003. Averaging kernels for DOAS total-column satellite retrievals, *Atmos. Chem. Phys., 3*, 1285-1291.

Hakami, A., Henze, D. K., Seinfeld, J. H., Singh, K., Sandu, A., Kim, S., Byun, D., Li Q. 2007. The adjoint of CMAQ, *Environmental Science and Technology, vol. 41, no. 22*, 7807–7817.

Sandu, A., Daescu, D., Carmichael, G. and Chai, T. 2005. Adjoint Sensitivity Analysis of Regional Air Quality Models, *J. of Computational Physics 204*, 222 – 252

Sandu, A., Daescu, D., Carmichael, G. 2003. Direct and adjoint sensitivity analysis of chemical kinetic systems with KPP: Part I—theory and software tools, *Atmospheric Environment 37*, 5083–5096

Zhu, C., Byrd, R. H and Nocedal, J. 1997. L-BFGS-B, FORTRAN routines for large scale bound constrained optimization, *ACM Transactions on Mathematical Software, Vol 23, No. 4*, 550-560.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

The role of predictive models in energy efficiency optimisation of industrial plants and buildings

Petr Stluka, Jiří Rojíček, Karel Mařík

Honeywell Prague Laboratory, V Parku 2326/18, 148 00 Prague 4, Czech Republic
Petr.Stluka; Jiri.Rojicek; Karel.Marik {@Honeywell.com}

Abstract: Current environmental, legislation, and economic conditions require industrial manufacturing and process plants to pay special attention to monitoring and optimisation of energy efficiency and carbon emissions. Also building owners are getting increasingly concerned by their energy bills and corresponding carbon footprint. Specific energy efficiency optimisation concepts have been developed and adopted over the recent years in both industrial and residential areas. One important aspect, which is common to both, covers the area of energy demand modelling and forecasting. The better models and forecasts can be used, the better energy efficiency measures can be implemented. The paper provides review of energy efficiency optimisation applications and respective energy demand modelling techniques.

Keywords: energy efficiency, monitoring, optimisation, predictive modelling, energy demand forecasting, weather forecasts, carbon emissions

1. INTRODUCTION

Energy efficiency optimisation concerns variety of different entities including industrial manufacturing and process plants, utility and power distribution companies, as well as individual buildings and houses (see Figure 1). These facilities can be categorized from various perspectives. Utilities and power generation companies represent the energy *supply side* of the whole energy supply chain, while the industrial plants, buildings and houses are the places where energy is consumed and so they represent the energy *demand side*. The major differences between individual sectors are given by their different business objectives, which have also the dominant impact on the energy efficiency optimisation techniques being adopted nowadays.

Utility plants, cogeneration plants, heating plants and power plants are completely on the supply side. Their primary business objective is to generate appropriate amount of energy (power, steam, heating or cooling) to meet demands of their customers, which can be in both industrial and residential areas. Since energy is directly the key business driver, the energy efficiency optimisation is also consistently addressed in such plants, most typically by a hierarchical control and optimisation system that operates in a closed loop.

Manufacturing and process plants represent primarily the demand side although many of them have their own energy generation facilities, which however can cover only a smaller part of the entire demand. In general, these plants need to produce a mix of products at appropriate volumes to meet demands of customers in the downstream industries. Energy usually is the second largest operational cost after the cost of raw materials. Although it is a common imperative for industrial plants to reduce their energy consumption, the energy is frequently not included in the optimisation criteria defined at plant, unit or process levels. These are fully concerned with the first priority – delivery of products – and it means that the energy efficiency optimisation is typically addressed by systematic monitoring of energy consumption and continuous improvement of operating practices.

Buildings and houses are on the demand side although they can use their own generation (CHP) or micro-generation whenever available. Business objectives of various buildings may differ significantly depending on their type – e.g. office, hospital, university, army base, shopping centre, department store etc. The common constraint for any energy efficiency optimisation in buildings is the comfort of building occupants, which must be always satisfied. The optimisation process then usually includes monitoring of energy use and continuous improvement of daily operation practices – Yee [2004]. One of the aims is to influence people's behaviour.

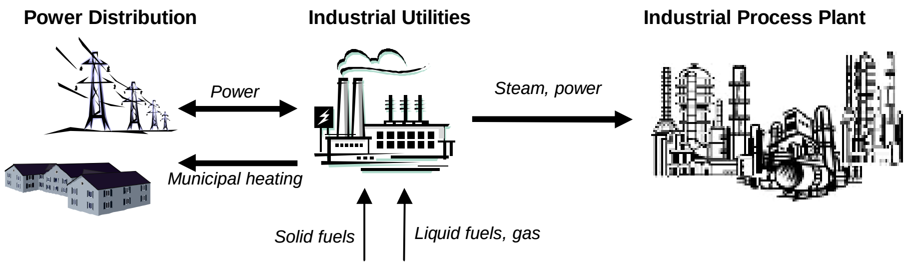


Figure 1. Major flows of energy between industrial plants, utilities and residential areas

The paper is organized as follows. *Section 2* introduces two different concepts of energy efficiency optimisation that are being implemented in various applications: continuous improvement and closed loop optimisation. *Section 3* then focuses on optimisation problems and respective predictive models used in the industry. This is followed by the same type of summary related to buildings. Lastly, the *section 4* concludes with a summary of commonly learned lessons.

2. ENERGY EFFICIENCY OPTIMISATION

Energy efficiency can be understood from two different viewpoints. Firstly, the system or process is operated energy efficiently if the amount of energy being used during the system operation is minimized. In this case the goal is to *minimize energy consumption* which is the extensive measure of energy efficiency. Secondly, the system or process is operated energy efficiently if its energy performance indicators are maximized - for instance, when the equipment like boiler, turbine, or heater is running at its peak efficiency. In this case the goal is to *maximize equipment efficiency* which is the intensive measure of energy efficiency.

Under specific conditions, the extensive or intensive energy efficiency measures can be directly controlled and then the natural choice would be to apply *closed loop optimisation* techniques. Otherwise the appropriate mechanism would be to systematically monitor those measures and aim at the *continuous improvement* of the overall process by eliminating causes of inefficient operation.

2.1 Continuous Improvement Cycle

The continuous improvement mechanism can be applied in cases when the energy consumption cannot be directly controlled, which is typical for the demand side facilities. The process follows four steps as illustrated in Figure 2, which correspond to the so-called Deming wheel, known from the field of quality management. The same high-level process can be applied to monitoring of energy efficiency in buildings – see e.g. Liu et al, [2002] – as well as in industrial plants.

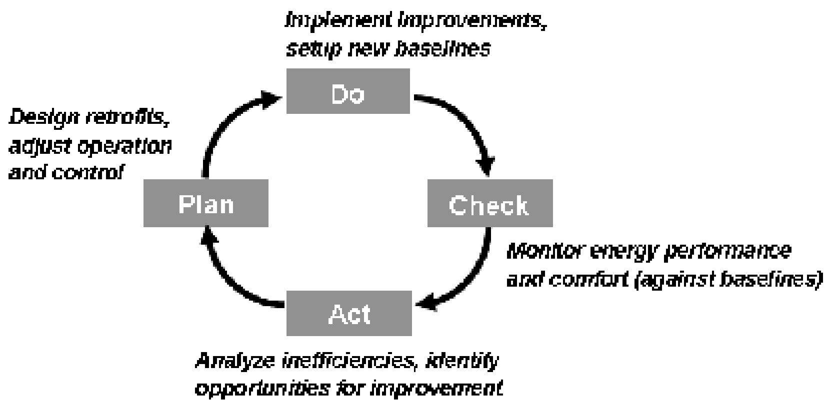


Figure 2. Continuous improvement cycle (Deming wheel) for energy efficiency optimisation

It starts from "*Do*" representing initial status of a system, e.g. a building after its commissioning or after completion of a significant retrofit. During this initial stage some ideal operation is usually achieved, which is then used as a baseline for further evaluation. The second step "*Check*" is directly about monitoring, which is implemented as a systematic comparison of the actual data (energy consumption, equipment efficiency) against the baselines determined in previous step. Outputs of the monitoring are generally deviations between the actual and ideal behaviour. The "*Act*" step is then focused on explanation of those deviations – in buildings the monitoring can be responsible for detecting symptoms of equipment faults, while the analysis focuses on their diagnosis. Results of the analysis are further used in the "*Plan*" step to determine actions that will lead to improvement of the current status, which can be also seen as elimination of existing operational problem. After these actions are implemented ("*Do*" step), the cycle starts again from its beginning. Application of this continuous improvement philosophy increases awareness of existing operational inefficiencies, which are then systematically eliminated.

2.2 Closed Loop Optimisation

Closed loop optimisation concepts are typically used in the energy generation plants (utility plants, combined heat power, etc.). The plant-wide optimisation may include applications organized in several layers as described by Havlena [2007].

The basic level is focused on *real-time optimisation* of individual pieces of equipment – basically the pressure control related devices like boilers, letdown valves and vents, but also other types of more complex equipment including turbo generators or condensing turbines. Multivariable predictive control techniques fit very well into this area as described e.g. by Findejs [2008].

The second level applications deal with the problem of *optimal allocation of load* between several pieces of equipment running in parallel. This task is usually executed in real-time to ensure fast response to dynamically changing conditions and requirements coming from the process plant.

Lastly, the third level applications optimise operation of the utility plant over significantly longer periods of time – ranging from hours to days – taking into account multiple possible configurations of the utility plant that can be selected for meeting the energy demand requirements. Flexible starts and stops of some pieces of equipment are assumed, which are mathematically translated into MILP type of *optimal scheduling* problem.

The concrete implementations may differ so that the second level applications can either be bundled with the real-time optimisation components into one solution package, or the load allocation problem can be solved as a subtask of the multi-period MILP optimiser. Solution details and practical results achieved when optimising operation of a CHP plant using the second approach were described by Schindler [2004] or more recently by Mařík et al. [2008]. Both approaches also differ in the way how they deal with varying energy demands. In the real-time optimisation approach this variable usually is considered as the disturbance variable that comes from the outside of the utility plant and in the relatively short time frame from 5 to 60 minutes they cannot be efficiently predicted.

On the other hand, the utility plant scheduling application cannot generate reliable schedules without having a reasonable energy demand predictions for the given time interval. Generally, those demand predictions can be combinations of two different types of demand: (a) *demand of the process plant* that is dictated by production objectives and needs in individual technological processes; (b) *demand of residential areas* that is driven by behavioural patterns of their inhabitants. Predictive models for both types of demand will be described in sections 3 a 4.

3. INDUSTRIAL PLANTS

Due to significant complexity of the production, a hierarchical top-down approach is adopted nowadays as the standard for managing industrial plants, which includes production planning, scheduling, and real-time optimisation layers. Outputs of corporate planning tools are fed into plant-wide production planners and schedulers that generate targets for individual processing units. Those are further projected into real-time optimisers (multi-unit or single-unit) in the form of limits. Given the overall complexity and primary focus on production, the currently used production planning and scheduling tools perform just material balance calculations without paying much attention to the energy intensity of the developed schedule. Energy costs are normally considered only at the process unit level.

Energy demand of the whole plant and its individual units is primarily given by the production plan or schedule that determines what products to make, by when, and how much. In a theory, there should be a good matching between such high-level production schedules and amounts of energy consumed on individual units. However, there typically are more or less significant deviations caused by drifting process conditions or disturbances like weather.

3.1 Predictive Models of Energy Use

Process unit operators have the responsibility for meeting production targets, formulated in terms of feed and product quantities and qualities, while minimizing operating costs on given unit. This is where the energy demand modelling comes into play. Predictive models of energy demand – covering electricity, steam, and fuel – are used for estimation of the future amounts of energy needed by major process units like atmospheric and vacuum distillation, or fluid catalytic cracking. The energy use is determined primarily based on the currently executed production schedule and its parameters, which can be:

- *Discrete variables* – modes, regimes, campaigns, types of raw materials, etc.
- *Continuous variables* – typically production volumes and qualities
- *Time variables* – start times and end times of planned modes and campaigns

The energy use is also influenced by changing key process variables and environmental conditions like ambient temperature, both typically not captured in the schedule.

Predictive models can be used for two different purposes: (1) generate estimates of expected energy consumption, which is consequently used as the operating target for unit operators; (2) advise operators on possible operational improvements in situations when the unit is deviating significantly from the expected energy use. Different level of detail is needed for each.

Energy targets = f (*Production targets, Disturbances*) (1)

Energy target models are not intended to provide precise numbers, and one of their important outcomes is the estimated variance of energy consumption, which indicates how consistently is the unit operated and controlled. Example of such a model is illustrated in Figure 3.

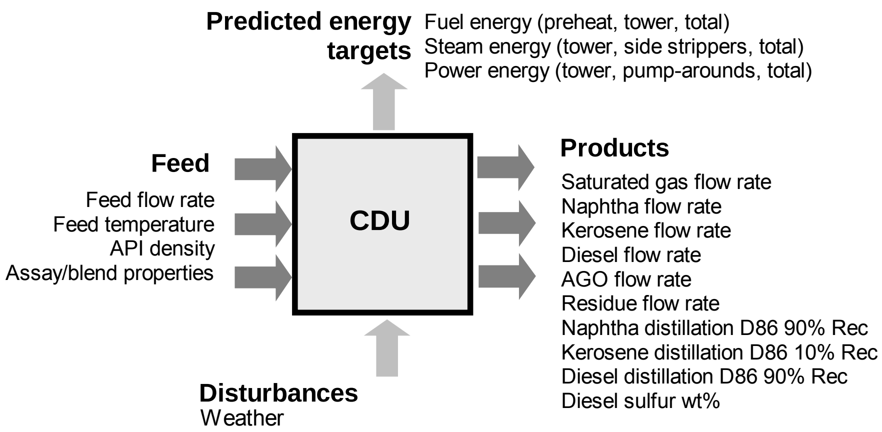


Figure 3. Variables used for energy target modelling on refinery crude distillation unit

Energy = f (*Production targets*, *Action variables*, *Disturbances*) (2)

Models used for advising operators must be based on more detailed modelling, requiring to capture relationships between production quantities, qualities, energy consumption and action variables.

The most popular modelling techniques used in the process industries include first principle models represented by various commercially available flowsheeting tools, and then also statistical data-driven modelling techniques, which build models by fitting relevant process data. The *first principle models* have the advantage of being based around fundamental understanding of the physics and chemistry of given process, and if accurate, can be used within a wide range of operating conditions. The downside is that these types of model are relatively costly to build and maintain, and sometimes also not easy to tune. The *data-driven models* can be derived by some form of statistical regression or non-linear fitting of data using various black-box modelling techniques. The data-driven models are more easily to build, they are also more robust and better dealing with natural variations in process measurements, but on the other hand, they are not useful outside the data set used to generate them.

When creating the high-level energy target models – formulated in equation (1) – for a process unit, the first choice usually is a simpler statistical model like the ordinary linear regression, which allows to calculate energy consumption based on actual production parameters. This approach works well until it is possible to effectively select input variables for the regression, using e.g. the domain knowledge about given process unit. When this is impossible, it is necessary to apply more sophisticated techniques for multi-dimensional problems:

- *Dimensionality reduction* – transformation of original variables into lower-dimensional set of new variables using e.g. Principal Component Analysis.
- *Multivariate statistical techniques* – e.g. linear or non-linear Partial Least Squares methods

- *Statistical learning techniques* – neural networks, regression and classification trees, ensemble methods like bagging or boosting.

Creating efficient detailed energy optimisation (i.e. operator advisory) models – equation (2) – is significantly more challenging because the set of considered variables is extended by the action variables. In this case the first principle models may represent viable alternative to the data-driven models. But also now it is necessary to consider overall effort needed for model setup and maintenance. An interesting strategy may be to use the first principle model for systematic evaluation of many possible process states. Running the model with changing input variables may help to generate a lookup table that can be subsequently used as a fast method for optimisation purposes.

Another approach may be represented by the local modelling techniques as described by Atkeson et al. [1997], which are based on the idea of using simple models around actual process states instead of complex global models. The local model is usually generated on-the-fly and it attempts to fit the training data only in the region around the given operating point. The local model itself cannot be used to find the global optimum values of action variables but it can advice about directions for possible improvement and this strategy can be applied recursively.

3.2 Predictive models of carbon emissions

Modelling of carbon emissions on individual processing units is frequently based on the concept of emission factors, which give reasonably accurate estimates of actual future emissions. Carbon emissions forecasting is applied to all major sources of emissions as described by Mondshine [2005]. For instance, in oil refineries these units include distillation units, conversion units, and hydrogen production units. The total forecast of future emissions is computed as a value aggregated over all sources of emissions, and is used as important input to utility optimisation. Future emissions can have impact on selection of fuels, as well as operating modes as discussed by Moore [2005].

4. BUILDINGS

Monitoring of energy efficiency in buildings can help building operators who are notified about fast changes in energy consumption or equipment performance, which allows them to take specific corrective actions – e.g. reconfiguration of the control system, or adjustment of equipment set points. The monitoring also provides great value to building owners, who may be warned about long-term trends and increasing wastage of energy. Consequently they can make more informed decisions on investments into modernization, system retrofits or equipment replacements.

4.1 Daily Profile Modelling

Modelling and forecasting of energy consumption in buildings is tailored to all specifics of this domain as much as possible. Because operation of any building follows more or less specific daily cycles, it is a common practice to describe energy consumption using the so-called daily profiles. The predicted daily profile for given day is then compared with actual energy consumption and any significant deviations are detected and appropriately addressed. Efficient modelling of daily profiles is an important pre-condition for advanced energy management applications.

Daily profile of a building reflects the way how the building is utilized. The variety of building types (e.g. administration, supermarket, casino, etc. as illustrated in Figure 4) is one of several challenges that need to be overcome to build efficient predictive model. The other challenges include:

- Energy consumption is influenced by changing *heating/cooling seasons*

- Buildings represent *dynamic environments* subjected to continuous changes, which may require relatively frequent adaptation of the model to new conditions
- Energy consumption depends on *stochastic factors* like is the number of occupants in given building, but also on *deterministic factors* like special days (holidays) when the building is closed.

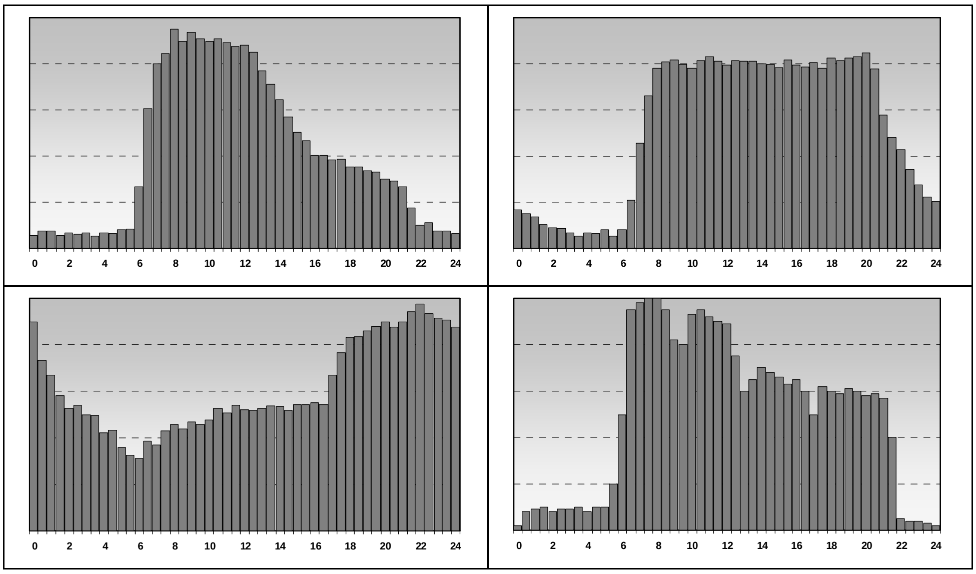


Figure 4. Typical daily profiles of electricity consumption for various types of buildings: administration, supermarket, casino, two-shift manufacturing

The state-of-the-art techniques for modelling of daily profiles can be classified as statistical regression modelling and first principle modelling. It has been recognized that statistical regression is able to cope with the typical modelling challenges quite efficiently, especially when being applied in an adaptive manner that assumes regular updates of model coefficients. Regression model can be developed even when some of the influencing factors are missing. Formally, the regression model has the standard form $y = \mathrm{f}(x_1,x_2,\ldots,x_N)$, where y is the energy consumption, which can alternatively be the fuel consumption or any other characteristic defined at the whole-building level, and $x_1,\ldots,x_N$ are independent influencing factors – mainly weather conditions, calendar-based variables, and seasonal effects as summarized by Beran et al. [2006]. Typically x_1 = time of day (hour), x_2 = type of day (working day, holiday, week-end, etc.), x_3 = ambient temperature, x_4 = occupancy level (percentage of building occupants compared to the normal status), etc.

Ambient temperature usually has the key impact, while the other environmental factors like humidity, wind speed, cloud cover, or sun irradiation can sometimes be used for better interpretation and finer modelling of the demand data. Calendar-based variables can efficiently help with capturing the behavioural patterns. These variables include time of day, which is defined on closed interval $<0;1>$ where 0 corresponds to 0:00 and 1 to the midnight, and also categorical variables like day of week, holiday and special day, which cause clustering of similar days into coherent groups.

4.2 Load Shedding in Buildings

Once a reliable model of daily consumption profile is available, it can also be used in predictive manner to calculate future energy consumption for one or more days ahead. The forecast horizon depends on the availability of future estimates of independent variables – for instance, the future ambient temperature can be available as part of public weather forecast. The main purpose of energy forecasting is to understand peak demands for particular days. Sometimes building owners need to make sure that their peak demand will not exceed its maximum level agreed with local electricity supplier. Exceeding the demand maximum automatically causes mandatory penalties that must be paid to the supplier. When being warned about such a risk a sufficient time ahead, the building owner can apply specific corrective actions that generally fall into category of “load shedding”. Its principle lies in reduction of the energy consumption during the peak demand hours that is achieved e.g. by switching selected non-critical devices off.

5. CONCLUSION

The paper summarized energy efficiency optimisation concepts used in the utility plants, industrial manufacturing and process plants, and in buildings. These include mechanisms of closed loop optimisation and continuous improvement, which can be effectively implemented if the future energy demand predictions are available. Related types of predictive models include:

- Models for *energy demand of residential areas* represent a well studied research field. The prediction accuracy can be very good because a relatively small number of influencing variables needs to be considered, and the daily consumption profiles are usually smooth.
- Models for *energy demand of individual process units* of an industrial plant represent much bigger practical challenge because specifics of every unit must be considered, always starting with a different set of variables, which is considerably larger than that used for energy demand modelling in buildings. The high-level models can be used for setting energy targets based on existing production plan, while the detailed models may be used to advise operators to take appropriate actions aiming at improved energy efficiency.

REFERENCES

Atkeson Ch.G., Moore A.W. and Schaal S. (1997), *Artificial Intelligence Review*, Volume 11 , Issue 1-5, pp. 11-73

Beran Z., Marik K., and Stluka P. (2006). Data-centric demand forecasting for utilities, In: *Proceedings of ESCAPE-16*, Garmisch-Partenkir-chen, 2006.

Findejs J., Havlena V. and Pachner D. (2008), Multivariable predictive circulating fluidized bed combustor control, *AT&P Journal* PLUS2

Haves, P.; Salsbury, T.; Claridge, T. and Liu, M. (2001). Use of whole building simulation in on-line performance assessment: modeling and implementation issues. In.: *Proceedings of Seventh IBPSA Conference*, Rio de Janeiro, Brazil, August 13-15, 2001.

Havlena V., Pekař J. (2007), Advanced Process Management for Power and Industrial Energy, In: ESCAPE-17, Bucharest.

Liu, M.; Clardge, D. E. and Turner, W. D. (2002). *Continuous Commissioning^{SM} Guidebook*. [on line] U.S. Department of Energy.

Mařík K., Schindler Z. and Stluka P. (2008), Decision support tools for advanced energy management, *Energy*, Volume 33, Issue 6, 858-873, Elsevier

Mondshine, M., and S. Messner (2005). Using forecasting tools to thrive in a carbon constrained world, *California Climate Action Registry Conference*.

Moore, I. (2005). Reducing CO_2 emissions. *PTQ Journal*, March 2005.

Petela, E. (2005). Managing energy. ***Process Engineering Journal***, May 2005, 17-18.

Schindler Z. Efficient load forecasting and optimal resource allocation, 2004, In: *Proceedings of AICARR international conference*, Milan.

Yee, G. and Webster, T. (2004). Review of Advanced Applications in Energy, Management, Control, and Information Systems [on line] Lawrence Berkeley National Laboratory, February 2004.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Cost Efficiency and Economies of Scale in the Supply Chain of Water for Industry in Thailand

Adsavakulchai S. [a] and Wachirathamrojn J. [b]

[a] *School of Engineering, University of the Thai Chamber of Commerce*
126/1 Vibphavadee Rangsit Rd., Dindaeng, Bangkok, THAILAND 10400
e-mail : suwannee_ads@utcc.ac.th
[b] *Ph.D. candidate in Logistics Management*
School of Engineering, University of the Thai Chamber of Commerce

Abstract: The availability of water resources in Thailand is important for agricultural and industrials. The economies efficiency in the supply chain of water for industry is based on the economic theory of demand and supply. Especially, the industry growth and the economic development are leading the water demand increase. The demand of water for industry illustrates the impact of infrastructures investment and water cost. In the economic perspective, cost efficiency and economies of scale in the water for industry can be estimated from the total cost of the resources supply and from the demand function. The main objective in this study is to analyze the applicability of econometric analysis to the supply chain evaluation of water for industry. Data for the study is obtained through the supply chain of water for industry. The preliminary result is to estimate a variable cost function using a model approach to assessing the water for industry in Thailand to improve its operating cost efficiency. On the going research, to develop a stochastic cost frontier model to measure the efficiency estimation is a stochastic frontier production function models

Keywords: Water for industry; Thailand; Stochastic cost frontier model.

1. INTRODUCTION

Eastern Water Resources Development and Management Public Company Limited was founded by the Provincial Waterworks Authority (PWA) on October 15, 1992 to be responsible for the development and management of main water pipeline systems in the Eastern Seaboard by supplying raw water to industrial estates, factories and waterworks authorities through the company's main raw water pipelines in Chonburi, Rayong, and Chachoengsao. The Company is principally engaged in the businesses of (a) development and management of the major water distribution pipeline systems in the eastern seaboard areas of Thailand and (b) procurement of raw water from government agency sources for commercial distribution to end users.

The company has been entrusted by the Royal Irrigation Department to construct an approximately 1.7 billion Baht water pipeline system linking the Prasae Reservoir and the Khlong Yai Reservoir. The project, completed in the beginning 2008 will serve to prevent raw water shortage and meet the increasing demand for raw water in the future. The construction of a reserve raw water reservoir in the Chachoengsao-Chonburi area with the capacity of 7.4 million cubic meters will also ensure service stability in the future.

In 2007, the Company and its subsidiaries recorded a total net profit of 440.68 million baht, representing a 13% decrease from 2006. Total revenue was 2,430.13 million Baht, increasing by 29.20 million Baht, or approximately 1.2%, from 2006. Revenue from each business is shown in Figure 1.

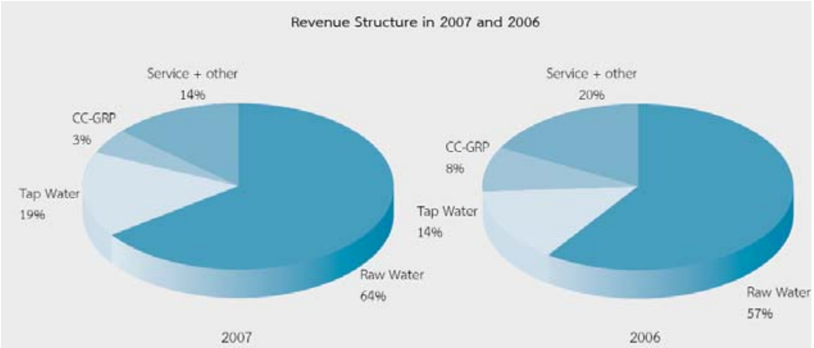


Figure 1: Revenue Structure in 2007 and 2006

The revenue from raw water business totaled 1,538 million baht, rising by 182 million Baht or 13% and corresponding to an increase of 5% in the quantity of raw water sold. Revenue from tap water business went up to 472 million baht, representing an increase of 127 million Baht or 37% as a result of sales to Ko Samui Waterworks, Sattahib Waterworks and Rayong Waterworks, and the increase in the number of new water users. Revenue from service dropped due to the implementation of the project to reduce losses of water in the service area in Region 2 of the Metropolitan Waterworks Authority. Revenue from sale of pipes was 63 million Baht, decreasing by 117 million Baht or 65% as a result of political instability. Big construction projects were negatively affected as the investment budget of the public sector for the year 2007 was delayed. The company's total operating expenditures were 1,676 million Baht, increasing by 58 million Baht or 4% from 2006.

To sum up, 2007 was a good year with plenty of water. The five main reservoirs (Nong Pla Lai, Dokkrai, Khlong Yai, Prasae, and Nong Kho) in the eastern region had a combined water volume of 450 million cubic meters (as of September 2007), or at 82.61% of their combined capacity. Thus, company will have enough reserve water for 2008 and in the future. Thus, the main objective in this study is to analyze the applicability of econometric analysis to the supply chain evaluation of water for industry.

2. METHODOLOGY

2. 1 Study supply chain of water

Raw water from natural resources is the upstream supply chain of water as shown in Figure 2. Mostly, no one think that it does not have a cost. Thus, the cost starts from manufacturing plant. From plant, water processing is to save in reservoir as warehouse before distribution. After that, the water will send to distribution center to distribute to industrial plant by plant.

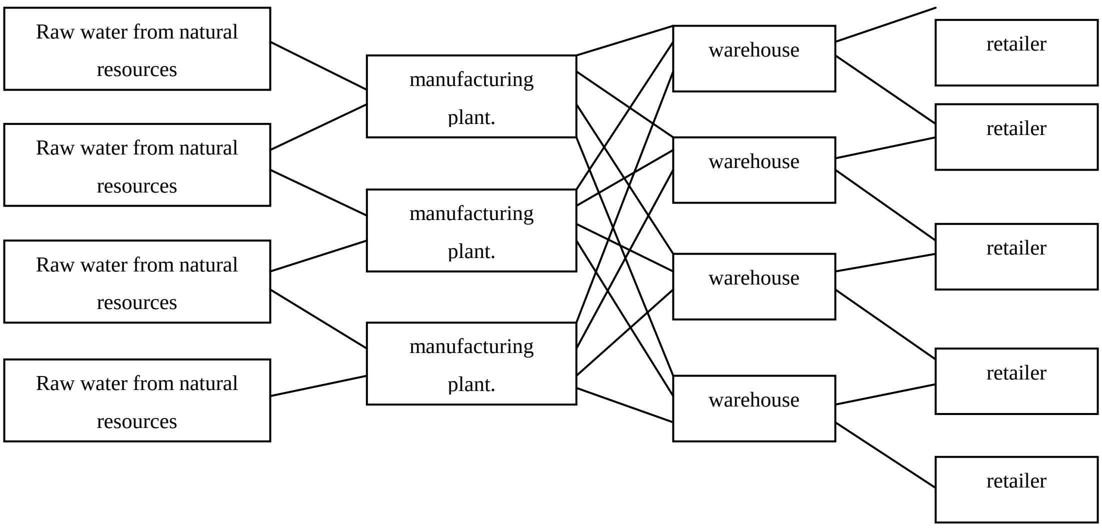


Figure 2: Supply chain of water

2.2 Logistics cost

In this study, logistics cost consists of four costs that are administrative cost, transportation cost, inventory cost and warehouse cost. There are several parameters related to each cost as following:

2.1.1 Administrative cost consists of labor cost, purchasing cost, and services cost.
2.1.2 Transportation cost consists of fuel cost.
2.1.3 Inventory cost consists of average inventory cost and interest rate.
2.1.4 Warehouse cost consists of depreciation cost of buildings and equipment and labor cost.

2.3 Logistics cost model

In this study, using SPSS is a tool for logistics cost model. Regression Analysis is one of the statistical analysis among parameters. Stepwise analysis is one of the statistical methods for parameter selection into the logistics cost model

3. RESULTS AND DISCUSSION

The results from regression analysis from each cost as following:

3.1 Administrative cost (AC) consists of labor cost, electricity, number of labor and office supplier. The results from regression analysis of administrative cost as following:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.99 (a)	1.000	1.000	.

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	7.109	6	1.185	.	.(a)
	Residual	.000	0	.		
	Total	7.109	6			

From the model, the parameters that related to administrative cost as shown in Eq. 1

$$AC = f\ (\text{labor, electricity, number of labor, office supplier}) \tag{1}$$

From equation 1, the correlation among parameters with coefficient of correlation (R= 0.99)

3.2 Transportation cost (TC) consists of labor cost, fuel cost, number of labor, interest rate and exchange rate.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.99(a)	0.99	1.000	.00627

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.842	5	.168	4287.70	.012(a)
	Residual	.000	1	.000		
	Total	.842	6			

From the model, the parameters that related to transportation cost as shown in Eq. 2

$$TC = f\ (\text{fuel cost}) \tag{2}$$

From equation 2, the correlation among parameters with coefficient of correlation (R= 0.99)

3.3 Inventory cost (IC) consists of average inventory cost and interest rate.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	1.000(a)	1.000	1.000	.

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.176	4	.044	.	.(a)
	Residual	.000	0	.		
	Total	.176	4			

From the model, the parameters that related to inventory cost as shown in Eq. 3.

$$IC = f\text{ (inventory cost, interest rate)} \qquad (3)$$

From equation 3, the correlation among parameters with coefficient of correlation (R = 1.0).

3.4 Warehouse cost (WC) consists of depreciation cost of buildings and equipment and labor cost

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.99 (a)	1.000	1.000	.

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.592	4	.148	.	.(a)
	Residual	.000	0	.		
	Total	.592	4			

From the model, the parameters that related to warehouse cost as shown in Eq. 4.

$$WC = f\text{ (depreciation cost of buildings and equipment, labor cost)} \qquad (4)$$

From equation 4, the correlation among parameters with coefficient of correlation (R = 1.0).

3.5 Total logistics cost (TLC)

From Eq. 1 to Eq. 4, the total logistics cost is the summary of each cost as shown in Eq. 5.

$$TLC = AC + TC + IC + WC \qquad (5)$$

To conclude, the results are the total logistics cost depend on administrative cost. This is because of the advertising and public relations expenses as shown in Figure 2.

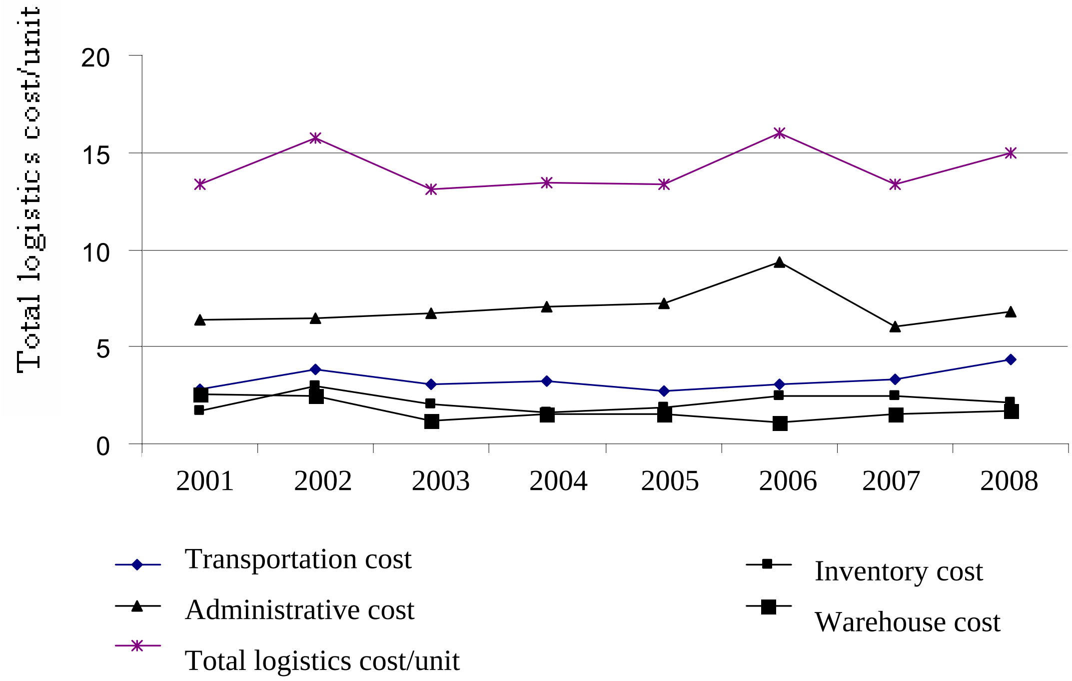


Figure 2: Total Logistics cost of water for industry

4. CONCLUSION AND RECOMMENDATIONS

Key risk factors that may affect cost efficiency and economies of scale in the supply chain of water for industry in Thailand include:

4.1 Investment in businesses that fail to achieve their goals. Over the years, the company has expanded its operations from raw water to also include tap water, drinking water, and pipe & equipment businesses. While the tap water business has enjoyed steady growth, the performance of drinking water and pipe & equipment was still far from satisfactory.

4.2 Conflicts with communities concerning the use of water resources. Most of the construction projects are major pipeline projects such as that from Prasae to Khlong Yai, one of the company's main raw water resources. Their implementation and the use of water resources in the company's projects may have led to conflicts with communities in the area. However, the company has coordinated with the Irrigation Department to open a dialogue and try to end the conflict with each community. Other activities included the educational project for farmers through a demonstration plot using the water dripping method and the organization of youth camps to raise awareness of participants on the value of water.

4.3 Inadequacy of water resources to meet water consumption demand. At the beginning of 2007 the water volume in the main reservoirs was 40 million cubic meter higher than in 2006. The projection in 2007 indicated that there would be sufficient water for consumption throughout 2007. However, to mitigate this risk, the company has implemented a monitoring system to be on the alert for droughts. Preventive measures were put in place and the water situation was monitored on a weekly basis. A project to increase raw water capacity was implemented, with the provision of earth pits to reserve water to be used during the dry season. At the time of this reporting, the water volume in the main reservoirs was 240 million cubic meters, or at 87% of the total capacity. The completion of the Prasae Reservoir-Khlong Yai pipeline project was completed and ready to divert water at the end of 2007. As a result, the company's reserve water increased and was adequate for current and future consumption. The risk of inadequacy water resources was then reduced to a minimal level.

4.4 Increased cost due to rising energy costs. Pumping water from newly developed water resources and the rising cost of energy due to a higher Ft rate resulted in higher costs than anticipated. Although the company could not control the cost of electricity which varied with the price of oil in the global market; it was able to manage by measures such as pumping water during the time that the cost was at the lowest and improving the pumps' efficiency. The company has also prepared an energy conservation plan and explored alternative energy sources such as water, the sun and the wind.

4.5 Policies of the government and related agencies. Uncertainty in the policies of government agencies could affect the company's operations. For example, there was a lack of clarity regarding the budget for the Prasae-. Khlong Yai Pipeline Project. The Irrigation Department was coordinating with the Ministry of Agriculture and Cooperatives to allocate a budget or 1,008 million Baht for this project. Still, the company has always been accorded full cooperation from the government agencies involved. Furthermore, the company's business is in line with a cabinet resolution to promote industrial investment and tourism in the Eastern Seaboard.

4.6 Damage to pipeline systems. The company's pipeline system may face the problem of corrosion or construction of other infrastructure in the areas where the pipelines are laid that may lead to breaking or leaking. However, the company has taken industrial all risks insurance to cover every pipeline and has installed the cathodic protection system to prevent or stop rusting or corrosion of iron pipes. In addition, more parallel pipes have been laid to meet the increasing demand for water, making it possible for interruption. The company to distribute water efficiently without although corporate risk is closely monitored by the Risk Management Committee, the company has set risk management action plans to be implemented by departments to maintain operational risks at an acceptable level. Risks are reviewed at the end of every year and not limited to those relating to damage. The review also includes consideration of other parameters tat can indicate probable risks under changing situations.

It can be concluded that the increase was due to the administrative cost that went up by 284 million Baht or 108% from the provision for diminution in value of inventories from the drinking water business and pipe business as well the advertising and public relations expenses. The cost of service business decreased, corresponding with the decrease in revenue from the area where the project to reduce water losses was implemented. Interest payable is increased by 99 million Baht or 106% due to the completion of Bangpakong-Chonburi project. On the going research, to improve develop logistic network that results in cost savings, convenient zonal distributions, and responsive supply chain management operations. To develop a stochastic cost frontier model to measure the efficiency estimation is a stochastic frontier production function models about a spatial distribution plan and route sequencing is developed for water distribution.

REFERENCES

Antonioli, B. and Filippini, M. (2001). The use of a variable cost function in the regulation of the Italian water industry. Utilities Policy 10, 181-187.

Chang S-Y, Brill E D, Jr., Hopkins L D, 1982, "Use of Mathematical Models to Generate Alternative Solutions to Water Resources Planning Problems", Water ResourcesResearch 18 58-64

Garcia, S. and Thomas, A. (2001). The Structure of Municipal Water Supply Costs: Application to a Panel of French Local Communities. Journal of Productivity Analysis 16, 5-29.

Kim, H.Y. and Clark, R.M. (1988). Economies of Scale and Scope in Water Supply. Regional Science and Urban Economics 18, 479-502.

Knaap G J, Ding C, Hopkins L D, 2001, "Managing Urban Growth for the Efficient Use of Public Infrastructure: Toward a Theory of Concurrency", International Regional Science Review 24 328-343

Montginoul, M. & Rieu, T. (2001). Implementation issues related to different water pricing reforms: application to the Charente river Basin in France and irrigation management agencies in Morocco. In: Conference Pricing water. Economics, Environment and Society. Sintra, 6 and 7 September 1999.

Varela-Ortega, C., Sumpsi, M. J., Garrido, A., Blanco, M. & Iglesias, E. (1998). Water pricing policies, public decision making and farmers' response: implications for water policy. Agricultural Economics 19 (1-2), 193-202

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Modelling of Floods: Climate Changes and Disasters Analysed by Information Technologies within a Framework of Recently Introduced Flood Directive of EU

E.Zeman, J. Spatka, V. Rara[a], M.B. Butts[b], J. Sirotek[c]

[a]*DHI a.s., Na Vrsich 5, Prague 100 00 Czech Republic*

[b]*DHI, Agern Alle 5, Hørsholm, DK 2970 DENMARK*

[c]*GEODIS BRNO, spol. s r.o., Lazaretní 11a, Brno 615 00 Czech Republic*

Abstract: Climate changes and Natural Disasters require the attention of experts seeking for new IT solutions, which could be used in a new concept of e-environment. Floodplain management at the catchment scale must address flood risk as part of a holistic water management strategy. In Europe, water managers must address the complementary requirements of the Water Framework Directive and the Floods Directive. This paper presents the development and application of an integrated hydrological modelling tool, based a combination of rainfall run off modelling technique and full hydrodynamic modelling where a coupled 1-D and 2-D hydraulic modelling capability has been developed to address the requirements of floodplain managers being responsible for implementation of EU Floods Directive. This new approach has been applied already in several places in EU and fully supports implementation of EU Flood directive included flood risk management aspects. Flood risk mapping tool was used after extreme floods in larger scale in the River Morava Basin, in the Czech Republic and also for flood risk mapping in the area of Prague, Capital of the Czech Republic. The set of tools and complete methodology were developed by the team of DHI key specialists using high resolution topographical data and archived maps and aerial photographs for model calibration. This tool can also address new problems such as urban flooding, assessment of proposed flood mitigation measures and new type of flood management master plans which assess the cross-border impacts of changes in water cycle and man-made structure impacts in whole basin. Flood mapping brings new inputs in prediction of potential flooded areas under various hydrological conditions, which is appreciated by rescuers and agencies responsible for emergency measures under the flood threats. The paper presents not only the description of the technology highlights but also the first results of the technology being used in the Jiu River Basin, Romania.

Key words: EU Flood directive; flood risk mapping; flood risk management; hydrodynamic modelling

1. INTRODUCTION

Effective floodplain management at the whole basin scale must address flood risk as part of a holistic water management strategy. From a European perspective water managers must address the complementary requirements of the Water Framework Directive (2000/60/EC) and the Floods Directive (2007/60/EC). The Water Framework Directive (WFD) sets a framework for comprehensive management of water resources in the European Community. Key objectives of the WFD are:

- to protect and enhance the status of aquatic ecosystems (and terrestrial ecosystems and wetlands directly dependent on aquatic ecosystems)
- to promote sustainable water use based on long-term protection of available water resources
- to provide for sufficient supply of good quality surface water and groundwater as need for sustainable, balanced and equitable water use

When new the Flood Directive is implemented, one has to be certain that the objectives of WFD are met, in other words the implementation should be a synthesis of both directives.

2. FLOOD DIRECTIVE 2007/60 EC

The economic losses caused by floods have increased significantly over recent decades in Europe and many other parts of the world. In order to reverse this trend, assessments of the flood risk have to be assessed in accordance with the recently accepted Directive 2007/66/EC ("EU Flood Directive").

A three-stage process has been defined in EU Flood Directive, which is obligatory for all EU member states from the year 2007:

- Preliminary Flood Risk Assessment,
- Flood Risks Mapping,
- Flood Risk Management Planning.

Based on a long-term experience in the field of sater and modelling, DHI experts have developed a powerful methodology for the implementation of EU Flood Directive, and tested it over the last decade on the pilot studies in various territories. One step leads to the next in a logical sequence, which guarantees successful fulfilment of EU directive requirements. This methodology is effectively supported by set of DHI specific SW tools, which were improved to meet client requirements. The methodology is actually applicable when the topographical and hydrological data sets are ready and suitable for modelling tasks including the accuracy limits and error limitations, which keeps the uncertainty of mapping results within reasonable margins and makes the final product feasible.

2.1 Preliminary Flood Risk Assessment:

By the year 2011 the identification of areas with significant flood risks or foreseeable future risks has to be done for each member state. It covers especially:

- Assessing the applicability of existing flood maps for the Directive,
- Analyses of all historical floods available in the respected domain (see Figure 2),
- Evaluating of flood events probability,
- Assessment of the impact of currently planned regional or urban development plans on flooding issues,

2.2 Flood Hazard and Flood Risk Maps

Based on the Preliminary Flood Risk Assessment all EU member states have to complete Flood Risk Maps before 2013 year. In general two sets of maps are required:

- Flood maps where the flood prone areas will be defined for flood with given probability of occurrence. Further depths (see Figure 1), stream velocities and erosion/deposit areas are also the complementary part of these Flood maps.
- Flood risk maps include numbers of potentially affected inhabitants; assessment of potential economical damages/losses and definitely also potential environment damages assessment in flood prone areas (see Figure 5).

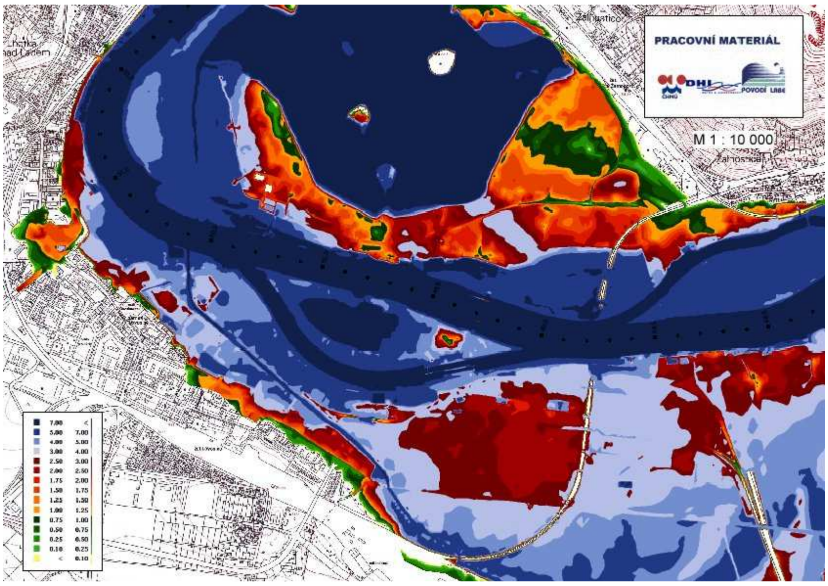


Figure 1 Example of map of depth

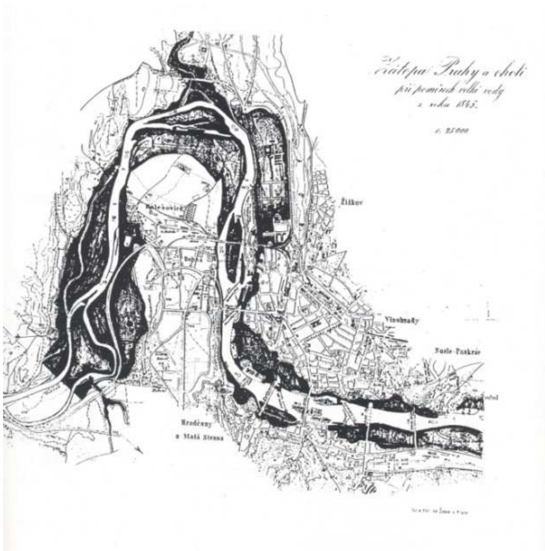


Figure 2 Historical flood map on Vltava river

2.3 Flood Risk Management Plans

Flood risk management planning represents the most important element of EU FD. This is a communicator and disseminator of the knowledge gained during two previous stages across the horizontal structures of governmental and non-governmental bodies dealing with flood protection, flood mitigation and flood struggle in general, included public involvement in this process. These plans must be finalised as the final round of the first planning cycle of EU Flood Directive by the end of the year 2015. Flood risk management plans mainly include proposals on how to reduce the losses of lives, property and environmental through flood prevention, protection of vulnerable areas and increased flood preparedness in each river basin. The way of processing of this flood risk management plans on IT platforms changes the information stream flow. Stake holders get the necessary information from servers or platforms and Water Management managers can disseminate the information in accordance with the agreed information channels.

Integrating powerful flood mapping capabilities with comprehensive hydrological modelling tool provides water managers with a valuable tool for both flood and Basin/Catchment management that permits a more holistic approach to flood risk as an integral part of more general water management issues.

2.4 Application of Flood Directive on Jiu basin

In 2006 the consortium of GEODIS BRNO (consortium leader), DHI (world-wide recognised model- provider) and CARTOTOP (local surveying office) won a large project for flood modelling in Romania.

The project is a part of an extensive programme of APELE Romane (Water Authority of Romania) to provide flood risk maps in all the main river basins across the entire country. The results of the studies and data are required by EU directives and have to be finished by 2013.

3. TOPOGRAPHICAL DATA REQUESTED FOR FLOOD MAPPING

With a high resolution topography data, it should be possible to simulate more accurately the detailed flow paths and properly capture the floodplain storage and this will lead to more reliable flood maps.

For the entire area of the river basin of the Jiu and other tributaries of the Danube, a hydrological model is being constructed based on DTM and LANDUSE data acquired with photogrammetric methods based on aerial images in scale 1:25000 (accuracy 1.5 m).

The domain is indicated for 984 km of the main rivers detailed in the topographical survey based on the photogrammetric evaluation of DTM (flight scale 1:4500, accuracy 15-20 cm), a classic topographical survey of the cross sections in 300 m distances along all the main rivers and bathymetrical measurements are provided. Mathematical hydrodynamic modelling will be made on the basis of this data.

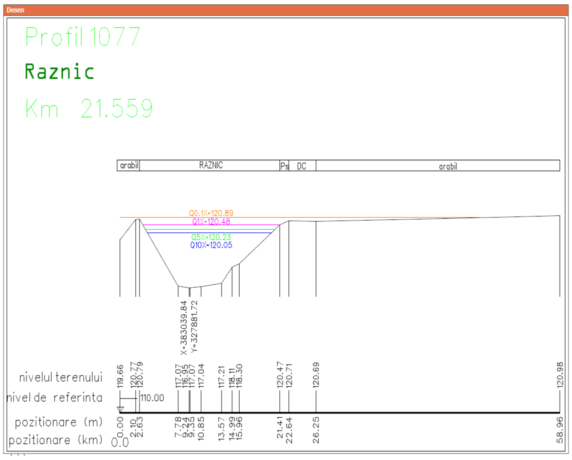


Figure 3 Water levels projected into Transversal section

Aerial images were triangulated using surveyed and signalized ground control points and parameters from D-GPS and IMU devices collecting data during the flight. A precise digital terrain model was measured using the most modern digital photogrammetric stations Z/I Image-station. Break-lines from 20 cm were measured and that the points in grid of 20 m. Accuracy of the DTM is 20 cm in altitude. Using DTM images were orthorectified and the final orthophoto-mosaic was created.

For each structure and cross sections on the river was established millage according to the river axis determined from the orthophoto-maps. The exact numbering of the structures was created. For each structures and cross section photo documentation was captured and detailed horizontal and vertical surveys were done. Based on approach a final longitudinal profile of the river was created.

Based on the orthophotomap the LANDUSE map was evaluated, which was later used in determining roughness characteristics and evaluation of potential of flood damages.

All the data will be delivered in Romanian coordinate system Stereo70 and height system Marea Negru 1975. All the topographical data will be delivered according to ANCPI standards.

4. MODELLING TOOLS AND THEIR INTEGRTAION FOR FLOOD MAPPING STUDIES

In Europe, water managers must address the key requirements of the Floods Directive (FD). MIKE family based coupled 1-D and 2-D hydraulic modelling has a capability to serve most of the requirements of the FD in a basic set-up.

On the Jiu River basin, the complete basin modelling is provided in this technology and appropriate modules of DHI SW are also a part of the delivery to Apele Romane.

The total catchment area is covered with hydrological model (precipitation-runoff model). For this model a description of catchment area (catchment area, catchment shape, catchment slope, Land use parameters …) was taken from digital elevation model of terrain. For Precipitation-runoff modelling, the whole Jiu River Basin was divided into number of sub-catchments draining specific areas and both, continuous MIKE NAM and event-based MIKE UHM hydrological models were created for each sub-catchment.

Flood areas along most important rivers, where the number of villages and other properties are located and where high probability of flood damage are observed, will be modelled using hydrodynamic models. Topology of these areas is described with much more precise geodetic data (high precision DEM). On the Jiu catchment area, approximately 970 km of rivers and their flood plains, were selected for hydrodynamic modelling

For hydraulic modelling three basic approaches are possible:

- 1D+ mathematical approach – MIKE 11, where the area of interest is schematized into mathematical model using set of cross-sections (taken from detailed digital model of terrain in carefully selected locations). Whole flooded area will be covered with looped network of river system and flood plain flows.
- 2D schematization – MIKE 21, which is based on computational grid placed on digital model of terrain and calculation in deep detail all hydraulic parameters in each computational grid. This procedure is very computation power demanding and is preferable in the areas of high interests, like urbanized areas of large villages, industry areas,…
- Combination of 1D and 2D approach – MIKE FLOOD, where the main river is solved by 1D schematization and flood plain is described by 2D schematization. This approach combines the advantages of both 1D and 2D principles in optimal proportions.

4.1 Practical outcomes of Flood Directive application

Above described methodology and data sources leads to generation of standard list of project results:

- Flood lines (see Figure 4)
- Water levels projected into longitudinal profiles
- Water levels projected into transversal profiles (see Figure 3)
- Risk maps according to EU Directive 2006/0005 (COD)
    - Potential flood damages
    - Number of flood exposed houses (people)
    - Danger zones
- Conceptual design of Flood protection measures
- Animations, Visualizations, HTML (see Figure 6)

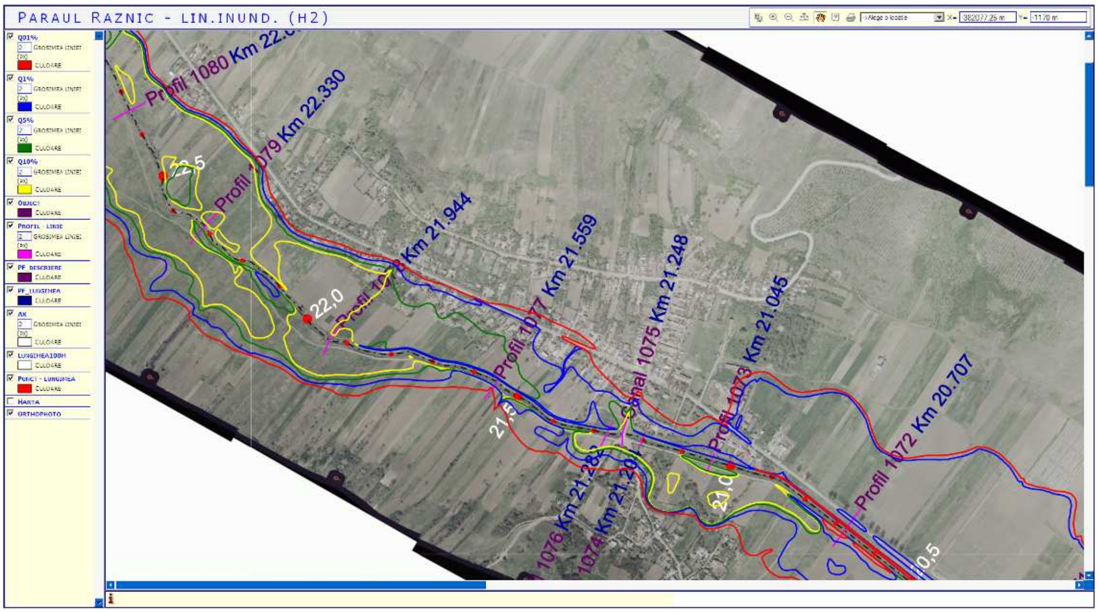


Figure 4 Example of flood line presented on background of Orthophoto map

In accordance with EU Flood directive 2006/0005, three type of hazard maps were generated using hydrological information (water levels, water depths and water velocities) taken from mathematical model, detailed terrain configuration taken from DEM and property type and position of buildings taken from Land Use maps.

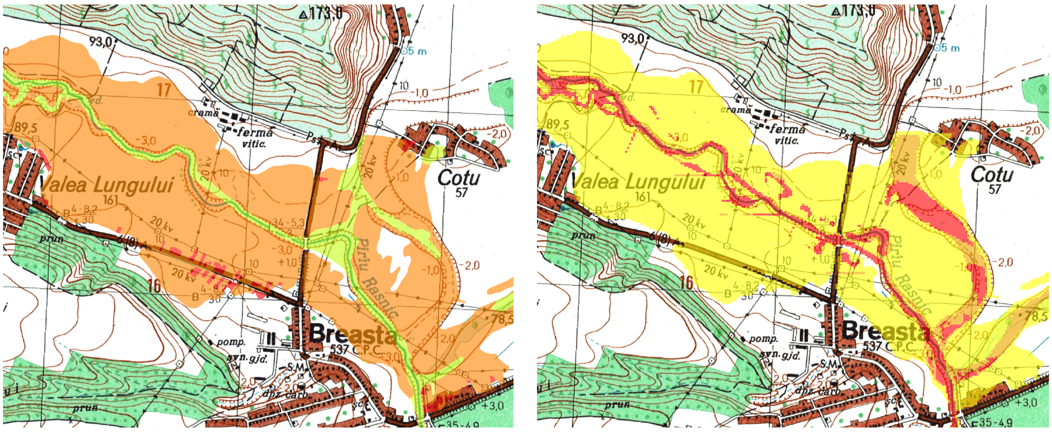


Figure 5 Left: Risk map – flood damages; **Right:** Danger zones

Risk maps of danger zones are based on combination of water level and water velocity parameters.

Map of flood damages was produced in relative scale, which, after evaluation of property values (prices), can be exchanged by currency scale.

Using flood maps and hazard maps, the potential protected areas can be identified, where flood protection measures would be proposed and evaluated.

The applied technology and methodology allow practitioners to produce also list of sophisticated project results, which would play a key role in the dissemination of project results:

- flooding animations
- self-navigating HTML project documentation, which allows to user easy preparation of different flood map composition and its printing in user selected scale.
- 3D visualization enabling virtual flights in the area of interest. User can view different layers for flooding lines and also has an access to technical documentation. This way of presentation can help to present project results to the broader public

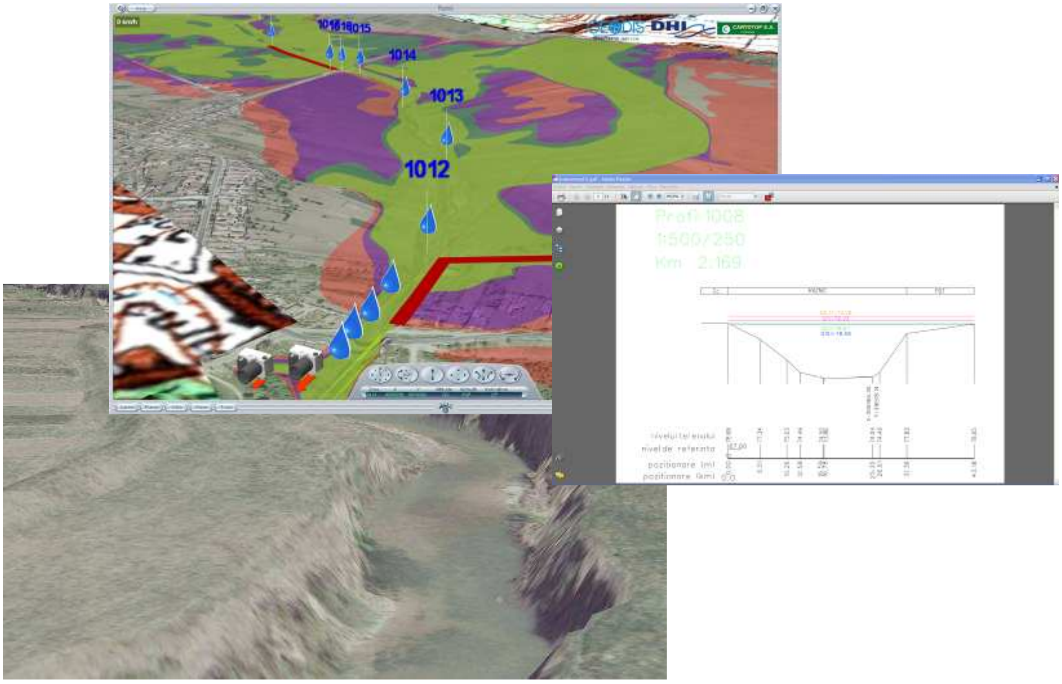


Figure 6 3D visualisation in GEOSHOW

4.2 Further utilization of obtained results

The selected methodology, which is being applied on the Jiu River basin, not only provides the direct results requested by FD in the format of different type of maps and technical drawings. One of the most important results for next extension of project results are the calibrated mathematical models of main river system. Such tools can be easily, and without significant extra effort and investments, used for detailed preparation of

- evacuation plans,
- emergency plans,
- operational flood maps for rescue services in the case of flood,
- flood forecasting and flood warning systems

5. CONCLUSIONS

The results of FD application on Jiu Catchment show, that the modelling approach, the data resolution and a proper hydraulic understanding of the floodplain flow are very important for correct application of EU Flood directives. The detailed results improve the understanding and general prediction of flooding on the major river system in Romania and assist the government to act (in time and space) in accordance with contingency planning based on flood risk management plans. The nature of the results enables the government or responsible ministry to disseminate results in modern technologies and this way to offer the information about evacuation plans, and flood risk management tools. Web based technologies and newly opened data stores might gain from the newly opened platform. The direction of information changes from sources to clients and clients get them in a new quality and much faster. Animation and various graphical outputs help understanding of non-experts as well.

Future development plans of regions and cities will get a proper guidance and platforms for future feasibility studies, which in urbanised areas must be provided in much detailed scale, however for these future studies, this project represents a master plan, which provides clear boundaries.

As a demonstration case could serve the effort of the Rumanian Government. It is a general belief that Romanian government, lead by Ministry of Environment have performed a very important and practical step for solving the task of FD of EU on very similar level of quality as other well established members of EU. The first completed section of the river to APELE and all delivered results, were accepted as being consistent with the TOR for the project as well as by the national technical commission. It is expected that during this year another couple of major rivers will be finished and within the next 2 years the whole project will be completed.

REFERENCES

Butts, M.B., E. Fontenot, M. Cavalli, C. Y. Pin, T.S. Jensen, T. Clausen, A. Taylor, **Development and evaluation of an integrated hydrological modelling tool for the Water Framework Directive and Floods Directive**, In Proceeding of International Conference FLOODrisk 2008, Oxford, United Kingdom, 30 September – 2 October 2008 (accepted).

Butts, M., J. Brandt, I. Cluckie, A. Dubicki, W. Szalinska, J. Špatka., „**Exploiting the potential of distributed modelling for flood forecasting and management**“, *Vodní Revue 8/2005*

Collection of case studies from UNESCO-IHE alumni, „**Water Education for a Changing World, Message from Alumni in the Field, Hydroinformatics tools saved lives during the floods in Central Europe in 2002, how theory and reality fit together**“, pgs.69-75, UNESCO-IHE 50 years of wise water, 2007, ISBN 978-90-73445-16-1

Sirotek, J., J. Špatka, **Flood management in Jiu River basin – the first river completed**, 2008 (will be published in GEODIS News)

Sklenar, P., Zeman, E., Spatka, J., Tacheci, P. „**Flood Modelling Koncept and Reality – August 2002 Flood in the Czech Republic**”, NATO ARW on Transboundary Floods: Reducing Risks and Enhancing Security through Improved Flood Management Planning, Oradea, Baile Felix, Romania, May 4-8, 2005, ISBN 1-4020-4901-3

Session 4

Global Monitoring for Environment and Security (potential for New EU countries for atmosphere and land monitoring)

Organized by Valere Mourtalier, Josef Aschbacher**, Ondřej Mirovský**, Jan Kolář, Tim Haigh, Jiří Hradec and Arno Kasch

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

*air*TEXT Air Quality Forecasting System

D. J. Carruthers and C.A. McHugh

Cambridge Environmental Research Consultants Ltd., Cambridge, England
(David.Carruthers@cerc.co.uk)

Abstract: *air*TEXT is an air quality forecasting and alert system for Greater London that has at its heart the ADMS-Urban dispersion model [CERC, 2008] and uses as input forecasts from meso-scale models that assimilate satellite data. The output is street-scale contours of pollution. By allowing for both the detailed effects of local traffic and the local effects of long range pollution, *air*TEXT forecasts air quality in individual streets and highlights the dramatic differences in air quality often found in different parts of the same city. *air*TEXT and the similar *Beijing*Air system are described in this paper.

***Keywords**:* Air quality; Forecast; Alert; Modelling.

1. *your*Air FORECASTING SYSTEM

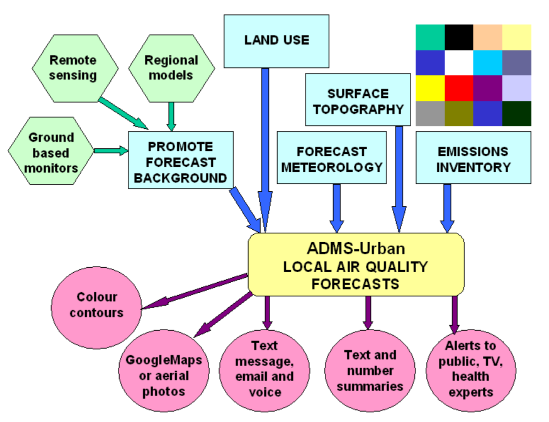


Figure 1. The *your*Air system.

*air*TEXT is the London version of CERC's *your*Air forecasting system that has the ADMS-Urban dispersion model at its heart and is driven by boundary conditions from mesoscale models such as Prev'air and EURAD run by European research organizations that assimilate satellite data to improve their model forecasts. ADMS-Urban is the most widely used advanced dispersion model for urban and regional air quality worldwide. It includes the Carbon Bond Mechanism (CBMIV) model of NO_X and VOC chemistry. The link between mesoscale models and ADMS-Urban was developed as part of the European Space Agency's PROMOTE project (www.gse-promote.org/). Figure 1 shows the *your*Air methodology. Sections 1.1 and 2 describe the system operation.

1.1 *air*TEXT Forecasts and Alerts for London

Launched in March 2007 the *air*TEXT system for Greater London and the Borough of Slough, a domain of 60km x 60km, generates street scale pollution contours twice a day that are posted to a web site (www.airtext.info) as well as sending pollution alerts to its 5,000 subscribers by text message, email and voice mail alert if pollution is predicted to be moderate, high or very high. The system uses regional background data from the Prev'air system, supplied by INERIS in France, which is derived using European Space Agency (ESA) satellite data as input to the model. Under the PROMOTE project an Integrated Air Quality Forecasting Platform (IAQ) comprising an ensemble of forecasts from regional European models is being developed and may be used by the *air*TEXT system in the future.

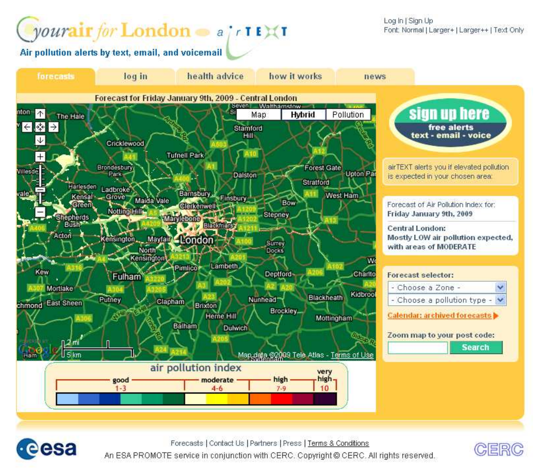


Figure 2. The *airTEXT* results on the web.

*air*TEXT alerts are triggered by pre-determined criteria e.g. concentration levels predicted at monitors or predicted areas of exceedence. Subscribers select to receive evening or morning alerts, the method of alerting and can manage their account for holiday periods. The alert accuracy of the system is around 70%. Comparing predicted concentrations with automatic monitoring data the agreement with monitored values (in terms of the COMEAP pollution band +/- 1 band [Department for Health, 2000]) is good: for ozone the accuracy across the sites is 83%-97%; for NO_2 73%-100%; for PM_{10} 42%-99%; and for SO_2 99%-100%.

2. FORECASTS FOR OTHER CITIES

*your*Air is being implemented for other cities, currently for Vienna and Vilnius and last summer the *Beijing*Air system was set up in a period of three months as a reliable, operational forecasting and alert system in time for the 2008 Beijing Olympics, for which air quality was a major concern. The model set up phase included the creation of a dual language web site with a sign-up facility for the alert service, creation of an emissions inventory, model set up and validation and establishment of the forecasting system with the live data feeds. The system was successfully installed and demonstrated at the offices of Beijing EPB and at Beijing Capital Normal University and training was also conducted at these institutes. The *Beijing*Air forecasts showed a good correlation with the measured Air Pollution Index values, good prediction of the mean values and a low normalised mean square error [CERC, 2009].

Figure 3. *Beijing*Air. Background concentrations from KNMI.

REFERENCES

CERC, ADMS-Urban User Guide, CERC, 3 Kings Parade, Cambridge, UK, CB2 1SJ, 2007.

CERC, Street Level Air Quality Forecasts for Beijing, Report to ESA, www.cerc.co.uk/services/Beijing_Report.pdf, 2009.

UK Department of Health, Air Pollution Bands www.advisorybodies.doh.gov.uk/comeap/statementsreports/apinfservice.htm, 2000

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I. Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Air pollution in Europe/CEE
History and outlook

Jaroslav Fiala
Czech Hydrometeorological Institute, Na Šabatce 17, Praha 4, Czech Republic
jar.fiala@chmi.cz

Abstract:

Pan-European reports stress that current levels of particles and ozone concentrations in western, central and southern Europe cause annually around 380,000 premature deaths, shortening average life expectancy almost one year and affect the healthy development of children. Around 47 % of EU's sensitive ecosystem area received nitrogen at levels that cause eutrophication and can lead to ecosystem alterations and biodiversity loss. In eastern and south-eastern Europe, the poor quality of the data precludes in-depth assessment of the state of air quality and its consequences. The limited available data indicate that the main health threat, as in the EU, is from small particles and their toxic constituents. In eastern Europe emissions of most air pollutants have increased by more than 10 % since 2000 as a result of economic recovery, increase in transport, and the persisting poor effectiveness of air pollution protection policies. Emission projections in eastern Europe for 2010 and 2020 point to a further increase, and greater efforts will be needed to achieve levels of air quality that do not give rise to significant threats to human health and the environment. In EU and the West Balkan countries, emissions of air pollutants are projected to decline during the next two decades as a result of progressive implementation of current and envisaged emission control legislation and continuing structural changes in the energy system. The projected emission reductions in EU and in West Balkan countries will reduce impacts on public health and ecosystems significantly by 2020, but not enough to ensure no significant threats to human health and the environment.

***Keywords**: PM_{10}, ozone, atmospheric emissions, PM_{10} and ozone ambient air concentrations.*

1. INTRODUCTION

Air quality assessment focussed on situation in 13 central and eastern European countries involved in the European Union's Phare Programme was published as EEA topic report No 16 (Fiala., J., Cernikovský, L., Kozakovič, L., Stedman, J.). It also provides information on emission sources and economic sectors contributing to air pollution, impacts on health and ecosystems as well as measures taken to reduce air pollution and its impacts.

Recent pan European assessment of air quality in Europe presented in the Air Quality chapter (J. Fiala et al., 2007) of the EEA State of the environment report No 1/2007 covers besides EU 15 countries and EFTA countries (Norway, Switzerland, Iceland and Liechtenstein) 12 new EU Member States (Bulgaria, the Czech Republic, Estonia, Lithuania, Latvia, Hungary, Poland, Slovenia, Slovakia and Rumania), the West Balkan countries (Albania, Bosnia and Herzegovina, Croatia, the Former Yugoslav Republic of Macedonia, Serbia and Montenegro), Turkey, also eastern European countries (Belarus, Republic of Moldova, the Russian Federation, Ukraine). It covers also the countries of central Asia (Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan and Uzbekistan) and the Caucasus countries (Armenia, Azerbaijan and Georgia).

The most recent assessments of the status and trends of air quality in the EEA region (EU 28 + EFTA + Turkey), based on the latest available validated air quality and air emissions data, are presented in the EEA and the EEA's European topic centre on air and climate change (ETC/ACC) reports (e.g. Mol, W. et al.: 2008).

Risk based assessment evaluating impact of air pollution on public health and the environment in EU was communicated in the CAFE Thematic strategy (European Commission, 2005). CAFE project results addressed and quantified health and environmental problems related to fine particles, ground-level ozone, acidification and eutrophication.

Air quality trends and status in the UNECE region is regularly presented in the reports of the EMEP programme (www.emep.int).

2. TRENDS OF ATMOSPHERIC EMISSIONS

Total PM_{10} emissions (emissions of primary particulate matter (PPM_{10}) and secondary particulate precursors (NO_x, SO_2 and NH_3) – Figure 1, have reduced by 44 % across the EEA-32 region between 1990 and 2006. Of this reduction, 43 % has taken place in the 'energy industries' sector due to the fuel-switching from coal to natural gas for electricity generation and improvements in the performance of pollution abatement equipment installed at industrial facilities.

Emissions of primary PM_{10} make only a small contribution to total particulate matter formation – 13 % in 2006. The majority of particulate matter is formed from emissions of the secondary particulate precursors. Of the particulate-forming pollutants, reductions of SO_2 that have taken place since 1990 have accounted for 60 % of the overall reduction in particulate emissions, with NO_x accounting for further 30 % of the observed reduction. The reduction in emissions of primary particles has accounted for only 6 % of the overall reduction.

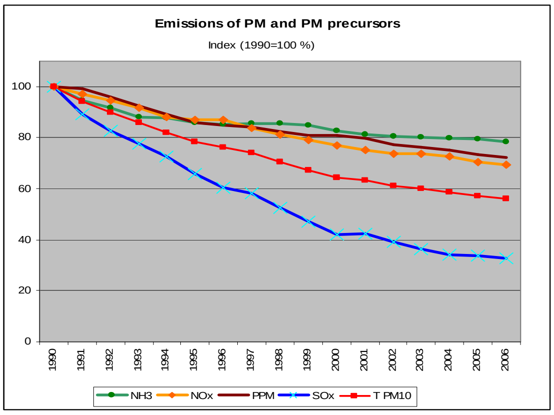


Figure 1. Emissions of primary particles and secondary particulate matter precursors (Index 1990=100 %). Source EEA.

Total EU-15 emissions of PM_{10} were reduced by 46 % between 1990 and 2006. EU-12 PM_{10} emissions were reduced by 54 % mainly due to reductions achieved in the industry and energy sectors which reduced their emissions by 70 % and 60 %, respectively.

The aggregated emissions of ground-level ozone precursor pollutants (NO_x, NMVOC, CO and CH_4) decreased by 37 % across the EEA-32 region between 1990 and 2006. This decrease has been achieved mainly as a result of the introduction of catalytic converters for vehicles and to a lesser extent by a switch from petrol-fuelled cars to more diesel cars. Together these changes have significantly reduced emissions of NO_x and CO from the road transport sector, the main source of ozone precursor emissions.

The European goal of achieving levels of air quality that do not damage people's health or the environment has still not been reached. EEA analysis suggests that 15 of the 27 EU Member States will miss one or more of their legally binding 2010 targets to reduce harmful air pollutants.

In eastern Europe (mainly in Moldova, Ukraine and the Russian Federation) emissions of most air pollutants have increased by more than 10 % since 2000 as a result of economic recovery, increase in transport, and the persisting poor effectiveness of air pollution protection policies.

3. AMBIENT AIR QUALITY

Despite apparent reduction of atmospheric emissions air pollution is still a serious environmental problem heaving substantial impacts on human health, ecosystems and crops.

Two pollutants, fine particulate matter and ground-level ozone, are now generally recognised as the most significant in terms of health impacts. Long-term and peak exposure can lead to a variety of health effects, ranging from minor irritation of the respiratory system to premature death.

A significant proportion of Europe's urban population still lives in cities where EU air quality limits, protecting human health, are regularly exceeded. According to the latest EEA data, since 1997 up to 45 % of Europe's urban population may have been exposed to concentrations of particulate matter above the EU limit set to protect human health. As much as 60 % of the urban population may have been exposed to levels of ozone that exceed the EU target value (Figure 2).

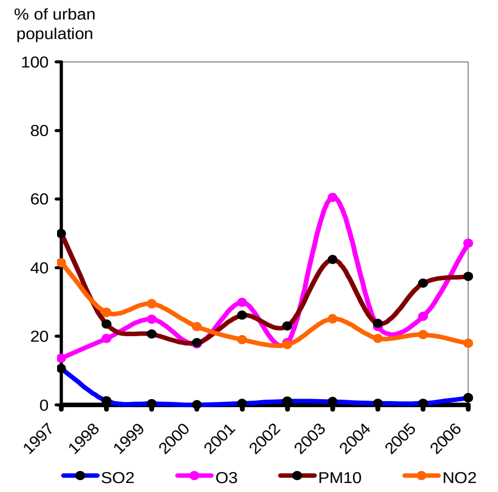


Figure 2. Percentages of the European urban population exposed to air pollution over the limit and target values. Source EEA.

While emissions of these two key air pollutants have dropped since 1997, measured concentrations in the air have remained largely the same. The reason of the discrepancy could be a combination of several factors: increased temperatures caused by climate change, hemispheric transport, influence of natural emissions of ozone forming substances released from nature and others.

As regards other regulated pollutants, the ambient air concentrations of SO_2 show a decreasing trend in line with emissions. Exceedances of the health related limit values are observed at a limited number of stations only. The emission of NO_x decreases but concentrations in urban and traffic areas are decreasing at a much lower rate. Compliance with the NO_2 limit value for annual mean concentrations is a serious problem in many urban and traffic areas. The ambient levels of CO as well as the benzene concentrations are in compliance with the limit values except for a limited number of traffic hotspot situations. The lead concentrations are well below the limit value for protection of human health. Several Member States have reported heavy metals (arsenic, cadmium, nickel) and benzo(a)pyrene regulated under the fourth Daughter Directive. The air pollution by these heavy metals is generally low: the concentrations are below the lower assessment threshold, for arsenic at the majority of the stations, and for cadmium and nickel at more than 90 % of the stations. The benzo(a)pyrene annual mean concentrations however exceed the target value at about 40 % of the 86 reported stations, namely in CEE (the Czech Republic). Further details on air pollution in the Czech Republic will be given in the oral presentation.

4. IMPACT OF AIR POLLUTION

The CAFE programme estimated a total of 348,000 premature deaths a year due to exposure to anthropogenic $PM_{2.5}$ emissions. At this level of exposure, average life expectancy is reduced by approximately one year. However, in the most affected areas of Benelux, northern Italy, and parts of Poland, the Czech Republic and Hungary, the average loss of life expectancy may reach two years.

In eastern and south-eastern Europe, the poor quality of the data precludes in-depth assessment of the state of air quality and its consequences. The limited available data

(Figure 3) indicate that the main health threat, as in the EU, is from small particles and their toxic constituents.

5. OUTLOOK

Following a CAFE analysis of a number of possible scenarios, the Commission adopted its Thematic Strategy on Air Pollution (European Commission, 2005) in September 2005. The strategy establishes interim environmental air quality objectives for the EU up to 2020. Results of the CAFE analysis are summarised in Table 1, which also shows the estimated benefits of the strategy.

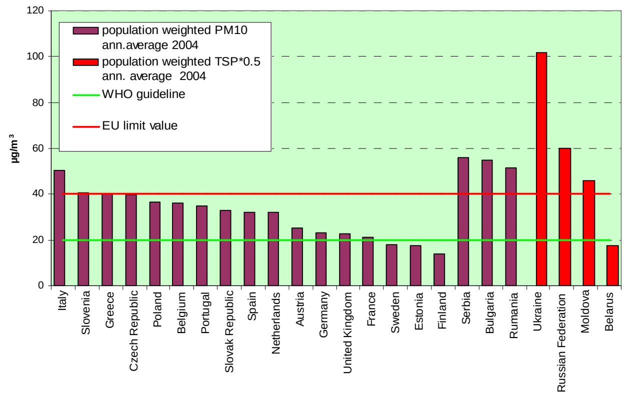


Figure 3. Average annual population-weighted PM_{10} urban concentrations in EU, SEE countries, and population-weighted PM_{10} concentrations based on monitored TSP concentrations in EE countries.

Table 1. Summary table of the CAFE analysis and the strategy (European Commission, 2005)

Level of ambition	Benefits					
	Human health			Natural environment (thousand km^2)		
	Monetised health benefits (Euro bn)	Life years lost due to fine particles ($PM_{2.5}$) (million)	Premature deaths due to fine particles and O_3	Acidification (forested area exceeded)	Eutrophication (ecosystem area exceeded)	Ozone (forest area exceeded)
2000	-	3.62	370 000	243	733	827
Baseline 2020	-	2.47	293 000	119	590	764
Thematic Strategy 2020	42–135	1.91	230 000	63	416	699
MTFR 2020	56–181	1.72	208 000	36	193	381

In line with the Thematic Strategy the European Commission plans to publish in 2009 a proposal to revise the current National Emissions Ceiling Directive (NECD), including stricter ceilings for the year 2020. National limits are likely to be proposed for fine particulate matter ($PM_{2.5}$) for the first time.

The NECD is mirrored by air quality directives setting limit and target values for major air pollutants. A new one called the Cleaner Air for Europe (CAFE) Directive was adopted in April 2008. For the first time it sets legally binding limit values for $PM_{2.5}$ concentrations to be attained in 2015. The European Commission is also taking countries to task for having missed earlier limits and, where sufficient measures have not been outlined to improve performance, has begun infringement proceedings.

REFERENCES

Fiala, J. et al: Air pollution chapter in the EEA State of the environment report No 1/2007, 'Europe's environment — The fourth assessment', (The Belgrade report)

Fiala, J., Cernikovský, L., Kozakovič, L., Stedman, J.: Air Quality in the Phare Countries in 1997, EEA Topic Report No 16/2002

Mol, W., van Hooydonk, P. and de Leeuw, F: European exchange of monitoring information and state of the air quality in 2006, ETC/ACC Technical Paper 2008/1

European Commission, 2005: Thematic Strategy on Air Pollution (2005). Communication from the Commission to the Council and the European Parliament. COM(2005)446

The European Meteorological Infrastructure - Key to Enabling a Range of GMES Core Services

EUMETNET [a]

[a] *EUMETNET Coordination Office c/o The Royal Meteorological Institute of Belgium, Ringlaan 3 Avenue Circulaire, 1180 Brussels, Belgium (adrian.broad@metoffice.gov.uk)*

1. EUROPEAN METEOROLOGICAL COMMUNITY (EMC)

The EMC consists of the National Meteorological (and Hydrological) Services (NM(H)Ss) of Europe (including some States that are not in the EU or EEA). In order to better fulfil their tasks and with the support of their governments, the NM(H)Ss have organised themselves at European level through 2 intergovernmental institutions:

- The European Centre for Medium Range Weather Forecasts (ECMWF) and
- The European Meteorological Satellite organisation (EUMETSAT)

To help the NM(H)Ss cooperate and coordinate their own activities, they have established a EUMETNET which, with its Members, operates a range of European wide activities and programmes.

2. EUROPEAN METEOROLOGICAL INFRASTRUCTURE (EMI)

The EMI can be considered to be the collective capabilities of the European Meteorological Community (EMC) needed to deliver their core services. The largest part of the EMI is the collective infrastructure operated at a national level by the NM(H)Ss.

The EMI is managed and operated by the EMC on behalf of the governments of Europe. It is estimated that, over the last 30 years, European taxpayers have invested in excess of 50 billion Euros in the EMI. It has taken decades to design and implement the EMI and it is still evolving. This represents a critical and available asset for an initiative such as GMES and a world class capability that should be exploited and not duplicated.

The capabilities and services operated by the meteorological community (in order to meet the needs of their national governments and customers) include:

- Gathering of European and global high quality, high frequency, real time, space and in situ observations of the weather and the oceans
- Creating, recovering and storing of high quality observations of climate for Europe
- World leading forecasting systems operating at a global, European and regional scale
- High performance computing
- Integrated, high availability, IT networks and databases
- High quality products and services delivered 24 hours a day, 7 days a week
- Highly skilled personnel
- Modern data access systems for users

All above mentioned are supported by quality assurance, standards, operational procedures and extensive R&D.

3. EUMETNET

The aim of the grouping is to enhance the collective and individual capability to serve environment management and climate monitoring and bring to national and European users high quality information and services. EUMETNET also coordinates external engagement from the NM(H)Ss, in particular with regards to the European Union. Networking allows the creation of cost-efficient EUMETNET Projects and Programmes that achieve objectives which could not be realised by NM(H)Ss working alone.

4. WHAT WE CAN OFFER

Essential to the success of the EMI are strong worldwide cooperation and networks (e.g. WMO, GEO,…). The organisations that operate the EMI have at their disposal important expertise related to GMES, that include for example:

- in-situ and space observations
- modelling and data assimilation
- providing operational 24/7 services
- developing and delivering versatile products and data to a diverse user community
- integrating data from several (also non-meteorological) sources into services for users.

The meteorological community that operates the EMI recognises the need for national capabilities to be integrated to deliver European level services and the relevance and importance behind the GMES programme. This community is collectively prepared to work with different EU agencies and partners to help exploit the EMI in delivering this key initiative. The meteorological community is working closely with the GMES Bureau to ensure the EMI can provide essential and cost effective support to a wide range of GMES services.

The EMI has been recognised as a partner of EC, ESA and EEA for the In Situ and Space Segment of GMES. The meteorological community, the EMI, and EUMETNET in particular, is a good example of European cooperation which acquired significant experience in the running of operational services and should therefore be considered as a reliable asset for EU-flagship GMES.

A testbed for INSPIRE and GMES services

V. Kraut, M. Schilcher, F. Straub

Technische Universität München, Fachgebiet Geoinformationssysteme Institut für Geodäsie, GIS und Landmanagement, Arcisstraße 21, 80333 München, Germany.

Abstract: The GMES program and the INSPIRE directive both have the goal to advance the knowledge about our environment. INSPIRE wants to achieve this by providing access to spatial data with environmental relevance collected from public authorities all over Europe. Therefore an European Spatial Data Infrastructure (ESDI) is built which enables standardised, web based access on this data. GMES on the other hand is building a network of in situ sensors and satellites which will be able to collect up to date information about various environmental topics. From this data map products are generated. Till now there exists no standardised way to access this data. This might be part of the problem why GMES is struggling to reach a broader user base. It seems reasonable to establish interconnections between these two components of the Shared Environmental information system (SEIS). An approach, taken in this project, is to provide GMES services as part of the ESDI. For the successful implementation of the ESDI a lot of tests are necessary. This is a result of new implementations of standardised interfaces in different software systems. Another challenge affects the institutional part of the implementation. INSPIRE requires authorities to provide data in a harmonized way. This often doesn't correspond with the way the data is gathered till now. The project INSPIRE-GMES-Testplattform wants to provide a testing environment where challenges which result from the implementation of the INSPIRE directive can be estimated and tested with the help of data from public authorities from different countries. This data is provided as INSPIRE services with the help of a broad range of software systems from different GIS vendors. The testing is done in a prototypical manner with the help of use cases which are oriented on actual administrative tasks.

Keywords: GMES; INSPIRE; OGC Geo Web Services; Interoperability; Testbed.

1. INTRODUCTION

With the rapid growth of the information technology also the way administrative tasks can be solved is revolted. Though web services have been present for several decades now - for example the Remote Procedure Call from 1976 - it took them some time to gain momentum. In the year 2000 they were celebrated as one of the most important trends in information technology.

So what exactly are web services? Though the term "internet" is often used as a synonym for "World Wide Web" (WWW), there exist many more ways to use this network. One of these is the web service technology. A very simple explanation would be that web services are for computers what web sites are for humans.

A more complex explanation would be that Web services provide a standardised way to access functions provided by servers. These services are described in a standardised machine readable way. The description is usually in form of a "Web Service Description Language" (WSDL) document. For the communication between web services and clients or other web services eXtensible Markup Language (XML) documents are sent in Simple Object Access Protocol (SOAP) envelopes over Hypertext Transfer Protocol (HTTP).

Web services are often used for the implementation of service-oriented architectures (SOA). A SOA is a design principle for the loose coupling of interoperable web services. It usually consists of users (or clients) who can access services which are provided by service providers. These services can be data, functions or platforms. They can be used to build business processes. Also services can be coupled with other services to solve more complex tasks. For the discovery of services the Universal Description Discovery and Integration (UDDI) registry is used [Sconnard 2002].

Since the year 2000 the trend of web services also reached the Geographic Information Systems (GIS) in form of services which were specified by the Open Geospatial Consortium (OGC). One of their first and still most popular specifications was the Web Map Service Standard (WMS) which enables the interoperable access to spatial data in the form of graphics (maps). Though geo web services are more or less similar to web services, there also exist some differences [Donaubauer 2004]. For example for a long time WSDL descriptions and SOAP envelopes were not used for geo web services. Fortunately nowadays the trend is to provide geo web services in a more akin way with common web services. This makes it easier to integrate them in business processes.

The geoinformatics sector always has been closely related to public administration. It seems only a logical consequence that approaches would be undertaken to use the benefits of geo web services to improve the way data from public authorities is accessed. On the one hand this trend has been promoted by eGovernment initiatives in several European countries. On the other hand up to 80% of the data which is gathered by public authorities has spatial relevance.

2. INSPIRE

At European level there have been several approaches to create a legislative justification for the use of web services for administrative tasks. This seems quite natural because European civil servants often face the problems which arise from the use of different ways to capture, store and provide spatial information.

These efforts finally were successful when in May 2007 the INSPIRE directive was published. Thereby the course for the creation of the European Spatial Data Infrastructure (ESDI) was set. This SDI intends to harmonize the access to spatial data all over Europe with the help of web services. Affected are all public authorities which are responsible for the capturing of data with relevance to the environment. Most of the governmental surveying authorities which provide general spatial data, needed for the proper use of the environmental data, are also affected.

The interoperability shall be achieved by the use of shared data models for the provision of the data. Furthermore five kinds of standardized Web Services to access and use this data and other web services are defined [Fichtinger et al 2008]:

- "discovery services making it possible to search for spatial data sets and services on the basis of the content of the corresponding metadata and to display the content of the metadata;
- view services making it possible, as a minimum, to display, navigate, zoom in/out, pan, or overlay viewable spatial data sets and to display legend information and any relevant content of metadata;
- download services, enabling copies of spatial data sets, or parts of such sets, to be downloaded and, where practicable, accessed directly;
- transformation services, enabling spatial data sets to be transformed with a view to achieving interoperability;
- services allowing spatial data services to be invoked." [EC 2007]

The INSPIRE directive tries to use existing standards as far as possible. Therefore a distinction is made between implementing rules and technical guidelines.

Implementing rules specify which functionalities an INSPIRE service has to provide. These rules have to be converted into national laws by the member states of the European Union and don´t contain details of the technical implementation.

These details are published in technical guidelines which aren´t part of the legislation process and thereby aren´t legally binding. This way it is possible to refer to existing technical solutions. It is also possible to react quickly on technological changes. As a result many of the INSPIRE services will be a subtotal of the OGC standards. This means, that not every OGC Service will be an INSPIRE service, but that OGC services which are conform to the requirements stated in the INSPIRE implementing rules and serve data corresponding to the INSPIRE data model will be INSPIRE services.

The INSPIRE services will be accessible over an INSPIRE portal which is provided by the European Commission and will serve as a central entry point to search for environmental data from public authorities all over Europe. The portal already is accessible under following URI: http://www.inspire-geoportal.eu/

Till now only a very limited set of functionalities is provided. A fully functional version is expected in the year 2010. Some of the services will be accessible without a fee, while others won´t. This depends on the fact whether an authority is responsible for cost recovery or not.

The ESDI is intended as an open infrastructure, which isn´t limited for services from public authorities. As long as other parties are able to provide services which fulfil to the specifications set out in the INSPIRE implementing rules, they also are allowed to be a part of the ESDI and to register in the INSPIRE Portal. This certainly makes sense for companies which generate and sell spatial data. It makes even more sense for an European initiative which intends to gather and publish environmental data on pan-European scale.

3. GMES

This European initiative is of course the Global Monitoring for Environment and Security (GMES) initiative. GMES is a joined undertaking of the European Commission and the European Space Agency (ESA) with the aim to provide information for environment and security observation.

This is achieved with the help of the components shown in the “GMES diamond” (Figure 1). The space and the in-situ systems provide data, which is delivered through services. For the smooth functioning of this workflow a complex process of data integration and information management is necessary.

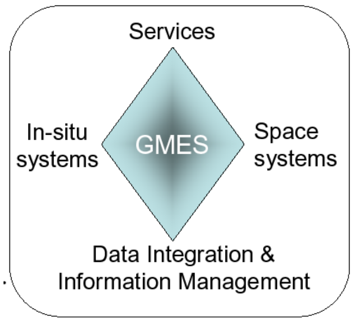


Figure 1 GMES diamond [EC 2004, p.8]

ESA is responsible for implementing the space segment of GMES. The space segment consists of existing national satellite missions of the member states (“contributing missions”). In addition five Sentinel missions with launch dates from 2011 to 2019 will be part of the GMES space infrastructure. ESA is also responsible for creating a framework which enables the standardized access to the earth observation data [ESA 2009].

The space segment is supplemented by the ground segment. This segment consists of a network of in-situ sensors. These sensors could for example be stationary sensors, airborne sensors or sensors on buoys in the ocean. Most of these sensors are operated by a wide variety of national authorities and research facilities and till now there exists no easy way to access this data.

This is where one of the main objectives of GMES kicks in. As stated for example in the GMES Action Plan (2004-2008) “GMES aims at coordinating existing as well as new technologies and systems to better meet a structured demand for information. […] To achieve this, GMES needs to make full use of data collected from space-borne, airborne and in-situ observation systems that is then delivered to service providers through an efficient data integration and information management capacity" [EC 2004, p.3]. GMES isn´t only about new satellite missions; the main objective is to bundle and generate synergies between the existing remote sensing activities in Europe, to close the gaps and to make better use of the gathered data.

To increase the usefulness of the data GMES wants to provide the data in form of services. GMES operates with a wider definition of services. In this case service only means a data product which is derived from the core services of GMES. User of these services might be government agencies as well as companies from the private sector.

GMES services can be classified in three major categories:

- Mapping: Generating a map product which covers in general a larger area of the earth surface and should be updated periodically.
- Emergency support: In case of natural hazards or other emergencies, GMES services can try to generate up to date data of the affected area.
- Forecasting: Forecasting includes a modelling process which generates short, medium or long-term predictions. Examples can be air quality or crop yields. [EC 2009]

Further it is separated between core services and downstream services:

- Core services: Core Services are standardized general purpose services which are regarded as an European public good. There are fast-track services for land management, the marine environment and services to aid emergency response. Further on two pilot service projects focus on security and atmospheric composition.
- Downstream services: Downstream services are derived from the core services and additional data to fulfil more specific data needs from users. They should be realized by companies and generate additional value to the core services.

A study for the extended impact assessment of INSPIRE stated already in 2003: "Where relevant, synergies with the GMES initiative will be sought in order to ensure coherence between INSPIRE and GMES" [EC 2003, p. 23]. And this is most certainly the case when it comes to enabling a shared way to access and use the data. Also in several GMES documents it is stated that it would be sensible to use the ESDI to provide GMES services [EC 2004, p 7; EC 2008, p 3].

Because of the fact that GMES is an initiative and no directive like INSPIRE it can´t specify binding rules for the provision of GMES services. None the less it is an important goal of GMES to create a possibility for standardised access to the services and a harmonized data structure. A project from the ESA, which deals with this issue, is the project Heterogeneous Missions Accessibility (HMA). It has the intention to build a Service Oriented Architecture (SOA) for interoperable data access on earth observation data. For the catalogue service the OGC CSW 2.0, ebRIM Profile was used, for online data access the OGC WMS and OGC Web Coverage Service (WCS) standards [ESA 2009B]. Also several other GMES projects already used OGC geo web services and considered the INSPIRE directive when it came to the provision of data and the design of interfaces. This can be seen in the SANY [SANY 2009] or the InterRisk project [InterRisk 2006].

4. INSPIRE-GMES-TESTBED

To be able to provide GMES services as part of the ESDI it is necessary to take the INSPIRE guidelines into account when developing GMES services. As a result there exists a demand for practical testing of the possibilities and requirements for this step.

INSPIRE itself raises the need for a testing environment. The reasons therefore are twofold. On the one hand INSPIRE influences the way the data has to be provided by defining new data models. Though it is intended to minimize the additional efforts for data acquisition, experience proved that it is a difficult process to provide data from different data models in a shared data model without losing too much content [GIS4EU 2008]. This isn´t only a technical challenge but also an administrative one, especially since there are no financial resources assigned to this task.

On the other hand INSPIRE doesn´t only point on existing standards for the provision of INSPIRE services but it also specifies certain abilities a service has to offer. For example in case of discovery services this would be the “Harvest” request which is not necessary for simple CSW services according to the OGC CSW 2.0.2 ISO Metadata Application Profile specification [OGC 2007, p.33] but which is required according to the INSPIRE Draft Implementing Rules for Discovery Services in form of a mandatory CollectMetadataRequest [EC 2008B, p.12]. Existing service implementations often don´t offer the full range of capabilities which are demanded for INSPIRE services. This is the reason why many GIS software companies have to alter their implementation of service interfaces to be able to provide INSPIRE services. These interfaces are based on complex specifications which easily can be misunderstood or in some cases even leave room for different interpretations. It takes some testing of the software interface till truly interoperable services are possible.

This said the Runder Tisch GIS e.V., a registered society with the aim to advance geoinformatics through cooperation and undertaking projects together with partners from science, economy and administration, wanted to provide a testbed for prototypical interoperability testing for its members. A project to implement this testbed was started together with 22 partners (Figure 2). These partners are national surveying authorities, GIS companies and service providers, universities and remote sensing companies.

Figure 2 Project partners of the testbed

The testbed is situated in the region of Lake Constance which stretches over three countries and thereby provides a heterogeneous GIS landscape (different GIS, data, data models, etc.). It consists of a central server with a collection of virtual machines as well as several distributed servers provided by project partners. This design allows on the one hand a flexible testing environment where all relevant parameters can be influenced; on the other hand the service architecture is closely related to reality. Through a security concept and agreements between the project partners a long-lasting operation can be guaranteed.

The testbed is built in several project phases beginning from March 2008 till July 2010 and will be operated for an even longer period. At the moment a service environment is developed which consist of OGC WMS, OGC WFS, OGC WCS and OGC CSW services as well as aggregate services which combine the afore mentioned services (Figure 3). The focus on OGC Web Services is due to the fact that the INSPIRE services will be heavily related to existing standards [EC 2008 C].

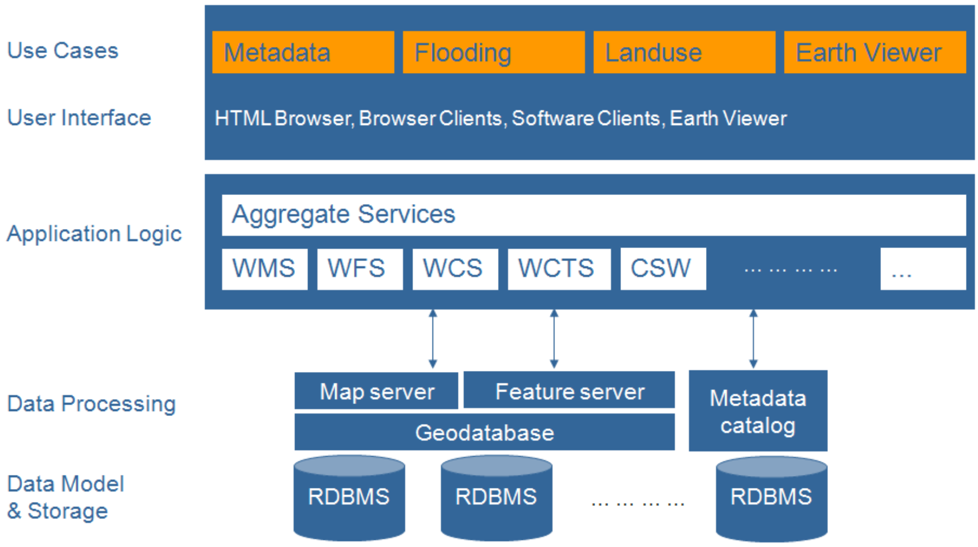


Figure 3 Conceptual architecture of the INSPIRE-GMES-Testbed

Phase II Metadata

In the current project phase, during June 2008 and February 2009, eight OGC CSW metadata services have been implemented with the help of project partners (Figure 4).

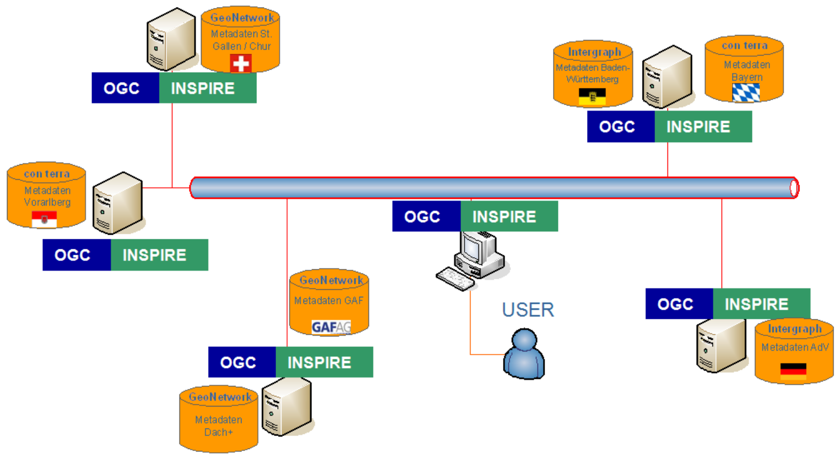


Figure 4 Conceptual architecture of phase II metadata

These services can be addressed with an aggregate service which alters the XML-results delivered by the services into a joined KML-file which can be displayed in Google Earth or Microsoft Virtual Earth to enable a different search approach for metadata (Figure 5).

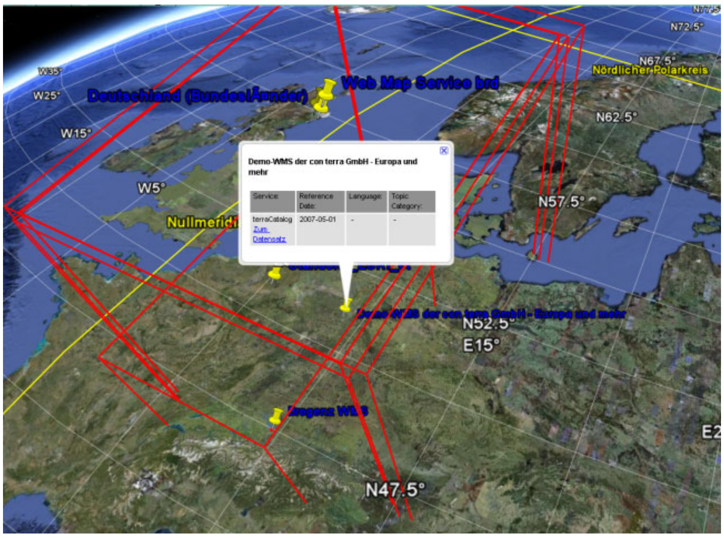


Figure 5 Results of the metadata search displayed in Google Earth

The experience gained in this project shows that there are still issues which have to be resolved for truly interoperable implementations of the OGC CSW standard.

Data issues:

- The datasets don´t contain all the elements required for ISO 19115 and INSPIRE conformance.
- Fields of the ISO 19115 Metadataset are filled in a heterogeneous way; specifications how the elements have to be filled aren´t always considered.
- The structure of the ISO 19139 XML files varies from catalogue system to catalogue system which makes it harder to process the file with external software.

Service issues:

- Most of the tested catalogue systems yet don´t offer all the functions which are demanded by the INSPIRE guideline.
- Especially functions which require the interaction of different catalogue systems, like Harvesting Requests, aren´t working properly till now. Reasons therefore are for example the different structure of SOAP envelopes generated by the catalogues or the different use of XML prefixes.
- Export of metadata sets and import in catalogue systems still proves difficult due to the different structure of the ISO 19139 XML files generated by the catalogues.

Phase III download and viewing services

In the next project phase download and viewing services will be implemented. This will be done by use cases with the topics flood simulation modelling and land use classes for regional planning. This project phase will be realized from 2009 to 1010 closely following the INSPIRE implementation process.

5. CONCLUSION

We hope we could show the expected benefits of an approach to join the efforts undertaken in the name of GMES and INSPIRE. The testing of interoperable services has to be guided by an official instance. In Germany for example this conformance testing is offered by the GDI-DE TEAM Engine [GDI-DE 2009], where catalogue systems can be tested against a central reference catalogue. It is none the less necessary to conduct practical tests of different systems in interaction because there are many issues which can´t be discovered by comparison against an archetypical catalogue. Often problems arise which simply can´t be foreseen. Practical tests can hardly be done in a centralized way because there is no legal

bond to participate in such tests. This kind of test can only be done on a voluntary base. The Runder Tisch GIS e.V. offers such a platform for the cooperation between different stakeholders in the GIS industry. With the help of the project partners from the Runder Tisch GIS e.V. such a testing environment could be created. The contribution of the partners is a central point for success. In this spirit we would like to thank the Runder Tisch GIS e.V. for the supporting and financing of the project and all the contributing partners for their effort, because a testbed can only be as good as it reflects the heterogeneous environment the implementation of the ESDI faces.

REFERENCES

Donaubauer, A. 2004: Interoperable Nutzung verteilter Geodatenbanken mittels standardisierter Geo Web Services. Dissertation am Fachgebiet Geoinformationssysteme der Technischen Universität München.

EC 2003: Contribution to the extended impact assessment of INSPIRE. URL: http://inspire.jrc.ec.europa.eu/reports/fds_report.pdf

EC 2004: COMMUNICATION FROM THE COMMISSION: Global Monitoring for Environment and Security (GMES): Establishing a GMES capacity by 2008 - (Action Plan (2004-2008)) URL: http://www.gmes.info/library/files/1.%20GMES%20Reference %20Documents/COM-2004-065.pdf

EC 2007: Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) URL: http://eur-lex.europa.eu/JOHtml.do?uri=OJ:L:2007:108:SOM:EN :HTML

EC 2008: COMMUNICATION FROM THE COMMISSION: Global Monitoring for Environment and Security (GMES): we care for a safer planet URL: http://www.gmes.info/library/index.php?action=standarddownload&filename=COM-2008-0748.pdf&directory=1.%20GMES%20Reference%20Documents&PHPSESSID =d12a1f25655287bc13a89f1d6a210b0e&PHPSESSID=d12a1f25655287bc13a89f1d6a2 10b0e

EC 2008 B: Draft Implementing Rules for Discovery Services (IR3) URL: http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/network/D3.7_IR3_Discovery _Services_v3.0.pdf

EC 2008 C: INSPIRE Network Services Architecture URL: http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/network/D3_5_INSPIRE_NS_ Architecture_v3-0.pdf

EC 2009: What services will be provided? URL: http://www.gmes.info/162.0.html

ESA 2007: Overall GMES Architecture URL: http://esamultimedia.esa.int/docs/ GMES/EC/2_coloc_GSEColoc_06Mar07_AP.pdf

ESA 2008: ESA starts to make data available for GMES. URL: http://www.esa.int/esaLP/SEM1Z9STGOF_LPgmes_0.html

ESA 2009: GMES Space Component. URL: http://www.esa.int/esaLP/ SEMOMS4KXMF_LPgmes_0.html

ESA 2009 B: Heterogeneous Missions Accessibility (HMA) Project Overview. URL: http://earth.esa.int/hma/

Fichtinger, A., Jaenicke, K., Kraut, V., Schilcher, M. 2008: INSPIRE für Entscheidungsträger. URL: http://www.rtg.bv.tum.de/content/view/262/50/

GDI-DE 2009: TEAM Engine. URL: http://www.gdi-de.org:8080/teamengine/

GIS4EU 2008: Newsletter - State of art and requirements for data harmonization, the GIS4EU contribution. URL: http://www.gisig.it/gis4eu/newsletters/GIS4EU_ news_3.pdf

InterRisk 2006: Pilot system. URL: http://interrisk.nersc.no/interrisk-pilot

OGC 2007: OpenGIS® Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile. URL: http://portal.opengeospatial.org/files/?artifact_id=21460

SANY 2009: SANY project skeleton. URL: http://www.sany-ip.eu/about_sany/project

Sconnard, A 2002: The Birth of Web Services, MSDN Magazine 10/2002 URL: http://msdn.microsoft.com/msdnmag/issues/02/10/XMLFiles

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

The HUMBOLDT project: Towards the harmonisation of spatial Information in Europe

Eva Klien[a], Christine Giger[b], Daniel Kristof[c], Thorsten Reitz[a], Santiago Caceres[d]
[a]*Fraunhofer Institute for Computer Graphics, A5 - Graphic Information Systems, Fraunhoferstr. 5, 64283 Darmstadt, Germany (eva.klien@igd.fraunhofer.de , thorsten.reitz@igd.fraunhofer.de)*

[b]*Swiss Federal Institute of Technology, ETHZ, Institute of Cartography, Wolfgang-Pauli-Strasse 15, CH-8093 Zurich, Switzerland (cgiger@ethz.ch)*

[c]*Institute of Geodesy, Cartography and Remote Sensing, Dept. GIS Developments, Bosnyák tér 5. H-1149 Budapest, Hungary (kristof.daniel@fomi.hu)*

[d]*ETRA Research and Development, Department of New Technologies, Tres Forques, no. 147, 46014 Valencia, Spain (scaceres.etra-id@grupoetra.com)*

Abstract: The four-year EU project HUMBOLDT (www.esdi-humboldt.eu, 6th Framework Programme) started in October 2006. This paper should not be considered as a scientific publication but as a rough summary of the contents of the HUMBOLDT project. HUMBOLDT aims at contributing to the implementation of a European Spatial Data Infrastructure (ESDI) that integrates the diversity of spatial data available for a multitude of European organisations. The main goal of the HUMBOLDT project is to facilitate the harmonisation of spatial information. The software tools and processes created will support the feasibility and advantages of an Infrastructure for Spatial Information in Europe as planned by the INSPIRE initiative, meeting the goals of Global Monitoring for Environment and Security (GMES). An essential element of the project is the development of user-driven, cross-border, GMES-related application scenarios in which the different components are applied and tested under realistic conditions. They will use data themes from the INSPIRE Annexes in their use cases and apply INSPIRE Implementing Rules where possible. Moreover, the scenarios are used to provide clear and realistic requirements on the development of software tools and components to support spatial data harmonisation processes. The first two years of HUMBOLDT already led to several major results. Based on the model-driven architecture and approach, a model editor was developed, which aims at the complete and easy specification of data schemas. In accordance to this approach, an alignment editor (tool for the specification of a transformation between source and target model) is currently under development. In addition, a framework was implemented, which supports the search for existing data and transformations according to given information product specifications.

***Keywords**:* Spatial data harmonisation; European Spatial Data Infrastructure; INSPIRE; GMES.

1. INTRODUCTION

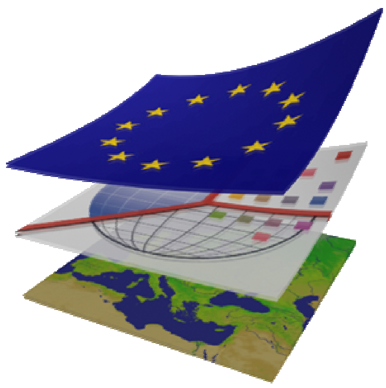
HUMBOLDT is a 6th Framework Programme EU project. Its project coordinator is the Fraunhofer Institute for Computer Graphics, Germany. HUMBOLDT started in October 2006 and will run for four years. In HUMBOLDT 28 partners from 14 European countries collaboratively work on the development of new tools and methods to support spatial data harmonisation processes. In HUMBOLDT harmonisation is understood as a process of transforming available source data into target data, according to given information product specifications, e.g. compliant to INSPIRE specifications for Annex I, II, III themes.

The project is funded by the GMES (Global Monitoring for Environment and Security) programme and aims at supporting the realisation of the European Directive INSPIRE in GMES-related applications. But HUMBOLDT is not limited to the goals of INSPIRE. Furthermore, the project follows the goal of being able to provide solutions that simplify the usage of spatial data almost wherever requested. Following Alexander von Humboldt, the project's philosophy is:

"Reuse the existing, extend by need, arrive at the ESDI and prosper by applications."

The project's name was selected with respect to Alexander v. Humboldt whose primary aim was to integrate the knowledge of his time to gain new insights and to further all areas of science. Following the same path, it is the aim of HUMBOLDT to base on the existing know-how within and outside GI community to manage and advance the implementation process of a European Spatial Data Infrastructure. This integrated network of systems providing data and services will allow the sustained use of existing services as well as the development of new applications and business models. The availability of data is – despite ongoing efforts – still highly scattered and heterogeneous. A harmonisation can contribute to the creation of new knowledge by combining data that could previously not be integrated at all - or only with prohibitively high effort. Also, entirely new processes that replace complicated existing activities and have a higher efficiency can be developed on the basis of the envisaged system.

1.1 Spatial Data Harmonisation in HUMBOLDT



In HUMBOLDT harmonisation is understood as a process transforming existing geodata into datasets that can easily be included into the usage processes at the data-requesting organisations. These transformations require as much knowledge as possible about the structure (syntax) and contents (semantics) of the involved spatial data. Some parts of that information remain in the heads of application experts but other significant parts can be formalized using so-called conceptual data models. Therefore, not only the format of the exchanged dataset has to be well-defined but also the conceptual model. The latter is providing sufficient information on how the real word is formally described in the computer system. For example, all data specifications for INSPIRE themes contain their respective conceptual models documented with UML (Unified Modeling Language www.uml.org), OCL (Object Constraint Language, www.omg.org) and some textual explanations. This technical approach using explicit representations of the conceptual data model in order to facilitate transformation processes is called "model-driven architecture" or "model-driven approach" (MDA, www.omg.org). The MDA is a major concept known from mainstream information technology, which serves as a basis for most technical developments in HUMBOLDT.

Knowing this, HUMBOLDT sees the need of using mainstream IT developments for data harmonisation and tailoring/customizing them to the spatial information community's needs. Only humans can handle several aspects of spatial data harmonisation processes, like identifying semantic similarities (e.g. "house" equals "building"). Other aspects can be carried out automatically as soon as the knowledge about the data models (both source and target) and the relations between them is available. HUMBOLDT mainly supports the elements of the harmonisation processes that can be specified formally and subsequently carried out by automated processes. Therefore, the HUMBOLDT framework and software tools deliver the necessary functionality to describe a desired information product (including conceptual data model(s)), search for existing data and mappings, or specify the mapping on the level of the data models and to transfer the data itself. A major requirement on HUMBOLDT is that the framework will allow the integration of the HUMBOLDT based solutions into several SDI architectures, to support for example both offline and online harmonisation processes.

2. ACTIVITIES AND ACHIEVED RESULTS

In its first two years, HUMBOLDT achieved many results with respect to different areas. This chapter contains summaries of the major results concerning scientific and technical aspects, introduced by some general remarks on the project's environment.

2.1 The Project's Environment

Main activities during the first year of the project were the implementation of an environment necessary to successfully carry out projects of the size of HUMBOLDT. This project environment covers the decision and management structures, the establishment of relationships to projects and initiatives relevant for HUMBOLDT, dissemination strategies and materials, and offers the ways of communication to experts and interested people in related fields. In the second year the environment was extensively used and extended due to the requirements, which occurred mainly during the first implementation phases for the technical components of the HUMBOLDT framework.

Due to the amount and heterogeneity of the partners' contribution to the HUMBOLDT project, special attention has been paid to leverage the knowledge across the partner structures to avoid misunderstandings and to ensure an efficient cooperation between the partners without loss of information. Therefore trainings for consortium members and externals have taken place. To ensure a fruitful interaction with other actors in the field of SDI development, HUMBOLDT has paid attention to ensure a broad visibility of HUMBOLDT within the GI community. In addition to several presentations at national and international events, publications in scientific and non-scientific journals, press releases, newsletters and on the website, HUMBOLDT has established contacts and an exchange of information with other projects such as BOSS4GMES, CASCADOSS, eSDI-Net+, RISE, MOTIIVE and further more.

Following the ideas of Alexander von Humboldt, the initial phase of the project has been used to look carefully into available technologies and concepts both from within the GI environment and from other fields. Furthermore the definition of potential HUMBOLDT users has been improved and the requirements resulting from this group have been selected. To create this overview of user needs, both the HUMBOLDT scenarios and public consultation have been used. Based on these overviews the important tool sets for HUMBOLDT have been identified and used to develop the concept of the technical prototype and the first version of the HUMBOLDT framework and tools. All the outcomes have been bundled into documents and other results. A list of the publicly available results can be found at the HUMBOLDT website http://www.esdi-humboldt.eu.

2.2 Scientific Results

In the first year of activities in HUMBOLDT was concentrated mainly on the state of the art in data harmonisation and management and related issues, such as the state of the art on harmonisations tools, software architectures, user needs and requirements. Based on the results of the analysis, the deficiencies of existing approaches were documented and a guideline for the scientific research to be carried out in HUMBOLDT was received. In the second year, the findings of the first year led to the development of new tools to support the data and harmonisation specifications processes. These tools (HUMBOLDT Model Editor, HUMBOLDT Alignment Editor) can be understood as facilitating the formalisation of application expert knowledge, necessary for the subsequent processing of existing data sets in order to carry out a harmonisation process.

The analysis of the collected documents during the state of the art analysis showed the broad range of the usage of geographic and spatial data, and consequently, the range in types of heterogeneity that occur in projects where information from different data providers is shared. Based on this still limited amount of documents a number of observations can be made.

The Ocean and atmospheric community depend on imagery and sensor data, and consequently here the focus of harmonisation activities is on metadata and metadata registries, standardised terminology, and a standardised reference grid, so that measured and forecast data from different data providers can be compared and linked.

The National Mapping Agencies on the other hand seem to be less involved in metadata (ISO 19115), although this is changing. Their focus is currently on the specification of common data models mainly for cadastre and official surveying purposes.

There is a broad usage and acknowledgement of metadata in the community providing environmental-related data. Nevertheless, although several data harmonisation activities are reported, there are only a few general approaches and reliable results.

An overview on specific harmonisation goals and purposes was developed, the possible solutions were presented, and it was indicated whether already an (open source) product that can be used, or a project that can serve as an example was found.

A major part of the investigations concerned the harmonisation of conceptual data models in order to facilitate the whole harmonisation process for spatial data. There are a number of steps, which have to be executed while fulfilling a data harmonisation task. They can be grouped into three major tasks/phases:

1. Specifying the harmonisation goals and approach (incl. creating a common target model)
2. Executing harmonisation process steps
3. Collecting feedback on the harmonisation result

All investigated methods and approaches showed that data harmonisation involves the knowledge of human beings. Only, after a first complete execution of a whole harmonisation process, some process steps or even the whole process can be automated.

This led to the following overall goal for the HUMBOLDT software developments (simplified) as illustrated in Figure 1.

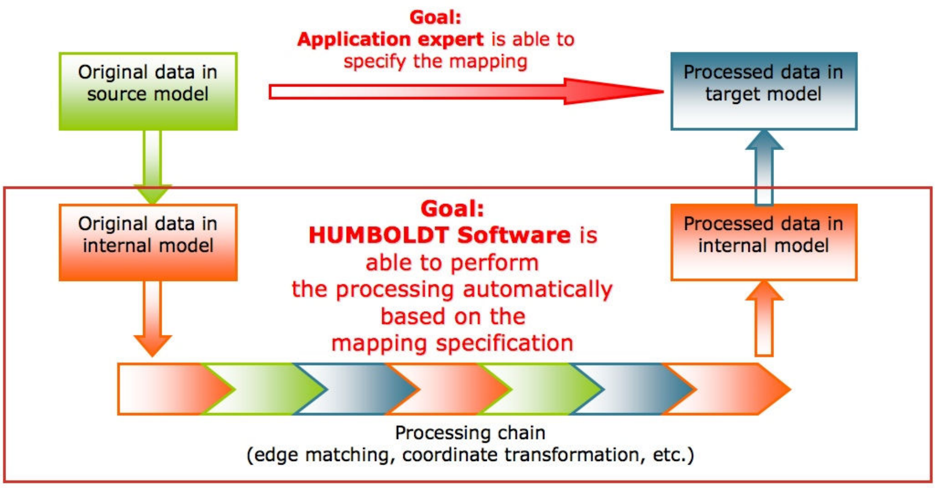


Figure 1: Help application experts to specify formally their target (desired information product) and the respective mapping process between source and target and let HUMBOLDT software subsequently execute the process as automated as possible.

These findings resulted in the development of the HUMBOLDT Model Editor, which is tailored to support the specification of conceptual models for spatial data (source and target) in the context of INSPIRE. Especially taking into account the major current results of the INSPIRE drafting team on data specifications. The current first prototype was developed mainly to support the further investigation of the harmonisation process itself but is now transferred from the scientific work packages of HUMBOLDT to the technical work packages in order to be tested and completed for actual use in HUMBOLDT.

In addition, the specification work on the HUMBOLDT Alignment Editor started in close cooperation with the technical work packages. This editor will support the specification of mapping steps based on given source and target spatial data models.

2.3 Technical Results

In this section the processes, methodologies, and results talking about HUMBOLDT from a technical point of view are briefly summarised.

A study about the state of the art of collaborative software development has been carried out. It provides an overview of current practices in international projects and developers communities for collaborative software development, with a special focus on the processes that aim at defining how distributed developers are brought to work concurrently toward a same goal.

The information is available in deliverable A4.1-D1 “State of the art in collaborative software development and definition processes”, which provides a general view of the domain of collaborative software development and helps to identify new practices and new ideas to manage, to process and to integrate the HUMBOLDT developments.

Furthermore, eight general specifications on eight important aspects of the process of collaborative software development for the HUMBOLDT Framework have been defined.

A set of documents has been created which serves as guidelines for developers. These are programming guidelines, development rules, a specification methodology and a document, which rules the building and testing of .Net and Java components.

The Prototype HUMBOLDT Framework specification has been developed following a system architecture approach to system design known as the Reference Model of Open Distributed Processing [ISO/IEC 10746]. The Reference Model for Open Distributed Processing (RM-ODP) is an international standard for designing open, distributed processing systems. It provides an overall conceptual framework for building distributed systems in an incremental manner.

RM-ODP defines standard concepts and terminology for open, distributed processing. In a generic way, the model identifies the top priorities for architectural specifications and provides a minimal set of requirements—plus an object model—to ensure system integrity.

Therefore the HUMBOLDT framework viewpoints based on the RM-ODP model are:

- *Enterprise viewpoint* of the HUMBOLDT framework. Information on typical business processes to be supported by the usage of the HUMBOLDT framework prototype and scenarios as well as use cases prepared for the HUMBOLDT prototype development.
- *Computational viewpoint* of the HUMBOLDT framework. This viewpoint shows the components of the framework, their collaboration and the interfaces used for that purpose (see Figure 2). It is described by the overall, logical architecture and the descriptions of individual components of the architecture as well as the collaboration of components.
- *Information viewpoint* of the HUMBOLDT framework. Data structures that are being used for storage and exchange between the services components defined in the computational viewpoint.
- *Engineering viewpoint* of the HUMBOLDT framework. It covers mainly the overall technical architecture of HUMBOLDT, the deployment and integration aspects of the framework like computers, networks etc.
- *Technology viewpoint* of the HUMBOLDT framework. Technology mappings for the logical and physical architectures of HUMBOLDT. The mapping of the logical components to physical components. The mapping of logical and physical components to the real libraries, protocols and other technical artefacts.

The current version of the HUMBOLDT Framework consists of the following components:

- Context Service
- Information Grounding Service
- Mediator Service Core Modules (Request Broker, Workflow service, Access Queue Manager and transformation Queue Manager)
- Context Service Web Frontend
- Standard GIS Client
- Model Repository
- Data Access Component Manager

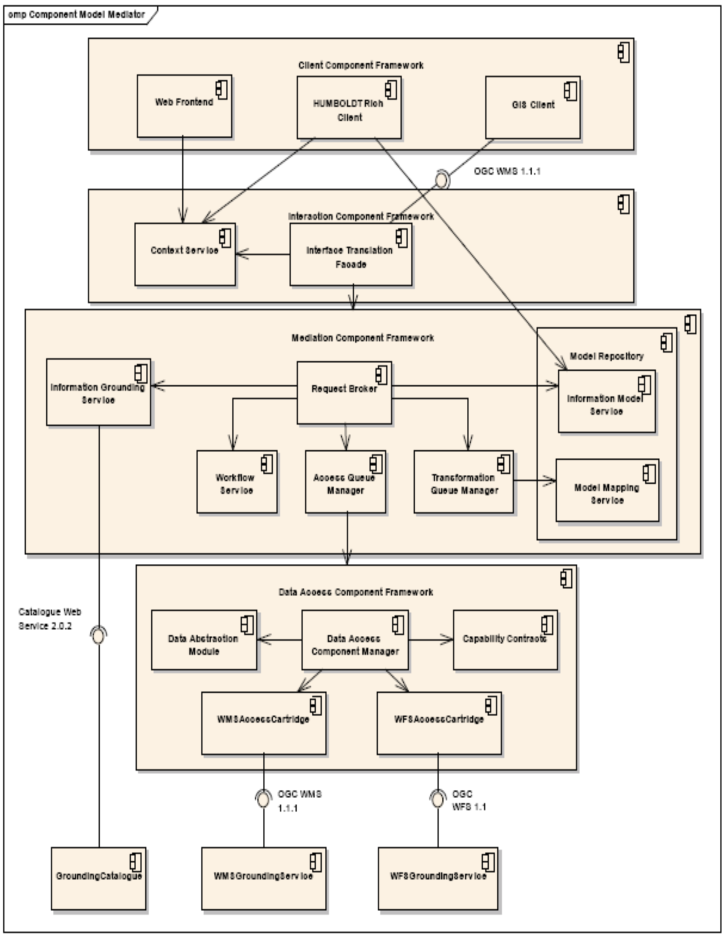


Figure 2: Component diagram showing the mapping of service components to virtual machines.

3. EXAMPLE FOR THE HUMBOLDT SCENARIOS

The HUMBOLDT Scenarios were selected to be realistic examples providing relevant user requirements as well as a suitable test bed for methods and tools developed in HUMBOLDT.

The HUMBOLDT scenarios are focussed on the following areas:

- Border Security
- Urban Planning
- Forest Management
- Protected Areas
- Flood Risk Management
- Trans-boundary Water Management
- Oil/Contaminants spill crisis impact and management for Oceans
- GALILEO Integration for Atmospheric Data Distribution

As an example for a typical use case in a HUMBOLDT scenario, we shortly describe the background and requirements in the Czech use case in the Urban Planning Scenario.

The use case in the scenario responds to the planning information needs arising from the deep administrative and political reforms that have taken place in the Czech Republic in the past 15 years, as well as the more recent impacts that followed the 1997 and 2002 flood disasters in the Czech Republic. The re-establishment of the municipality system, forming more than 6000 local authorities, 205 new districts and 14 regions was been supported by new legal provisions for the planning process based on the 1976 law, and subsequently replaced by a new spatial planning law from 2007.

The new law has created new rules for spatial planning influencing planning data usage, whilst the 1997 and 2002 flood disasters in the Czech Republic, have further redefined the key tasks of spatial planning and data needs placing greater emphasis on the management of natural threats and preventative measures. Increasing attention is being paid to the Integrated Emergency System, which integrates firemen, police, health rescue service etc. and complements existing spatial planning documentation. In the context of hazard assessment, spatial planning is now being coordinated with the Ministries of Agriculture (water management), Environment (water protection) and of the Interior (Firemen), and new types of dissemination tools to ensure public awareness.

In the Czech Republic the appropriate authority had to finish the first version of Territorially Analytic Backgrounds (TAB, in Czech UAP) by the end of the year 2008. Territorially Analytic Backgrounds contain findings and evaluation of status and development of territory by reasons of public policy, sustainable growth and change monitoring (Act 183/2006). TAB must be continuously updated. The complete updating must be done every two years. In the Czech Republic there are 219 responsible institutions (regions and some municipalities). Therefore, activities connected with acquisition, processing and providing of urban planning data in the Czech Republic are currently fragmented.

Concerning data and metadata management, there are several deficiencies and harmonisation tasks occurring in the current situation. Among them are:

- data sets with no known origin
- no clear data properties (e.g. data models, data formats, portrayal rules etc.)
- most of municipalities use proprietary closed desktop solutions
- almost no metadata at the level of municipalities
- no interoperable metadata solutions
- no metadata standards supported

The needed functionality, which has to be supported by harmonised data, would be:

- Search incoming information necessary for multi-criteria decisions; validation description, and preparation of metadata to assist in publishing the information;
- Initial virtual brainstorming (multi-participant) using analytical and decision support tools to set up and apply the decision model;
- Modelling of alternative scenarios using a combination of GIS (commercial and open source) and CDP in a common workspace;
- Discussions to validate the different scenarios and models and to agree on commonly acceptable solutions;
- Finalisation and visualisation of accepted plan.

ACKNOWLEDGMENTS

The work presented in this paper is the joint effort of the HUMBOLDT consortium; the authors listed above are the members of the Executive Board of the project.
The description of the entire consortium can be found at the project website at:
http://www.esdi-humboldt.eu/home/project_partners.html.

The HUMBOLDT project is financed under the Sixth Framework Programme of the European Union (Aerospace / GMES) for the period 01/10/2006 – 01/10/2010, with contract number SFP5-CT-2006-030962.

Www.airqualitynow.eu, a common website and air quality indices to compare cities across Europe

Karine Léger[a], Sef van den Elshout[b], Hermann Heich [c] and Fabio Nussio[d]

[a]: AIRPARIF, Paris, France (karine.leger@airparif.asso.fr)

[b] : DCMR EPA Rijnmond, Schiedam, The Netherlands (jel@dcmr.nl)

[c] : Heich Consult, Huerth, Germany (heich@heich-consult.de)

[d]: ATAC Spa,, Rome, Italy (fabio.nussio@atac.roma.it)

Abstract: Air quality is a public concern. This is partly due to the "right to know" principle embodied in European legislation. Despite this common legislation, the way air quality is being interpreted and communicated differs considerably. For specialists raw monitoring data for Europe are available but these are not usable by the general public. Easy to understand and internationally comparable air quality information from one city to another is scarce: there are almost as many air quality indices as air quality monitoring networks. The CITEAIR II project (Common information to European Air, INTERREG IVc) facilitates the comparison of urban air quality in near real-time by introducing two products: common air quality indices (hourly, daily and annual) implemented on a common website www.airqualitynow.eu using readily available simple IT-solutions. This paper describes those tools which both aimed at presenting the air quality of the participating cities in a comparable way and not to replace more targeted local information. Their value added is to provide, for the first time, a European and comparable picture of the air quality in near real-time easily accessible through a common platform and presentation of the results. The website is designed to receive and display data from each city wanting to join. The main part is dedicated to compare the cities index values using different time scales (hourly, daily or annual) and for two types of exposure thanks to a background and a traffic index. In addition, space is offered to cities for presenting themselves according to a common template, providing background information on their specific air pollution situation and associated reduction measures. Participating is easy: cities upload their data through ftp and the indices calculations are automatically made. The website provides a dynamic picture of the air quality and is updated each hour enticing viewers to make repeated visits. However, participation with only a daily update or with yearly data is feasible as well.

***Keywords**:* air quality index, public information, website, European platform.

1. INTERNATIONAL COMPARISON OF CITIES AIR QUALITY IN THE INTERNET

1.1 Introduction

The Framework directive and associated daughter directives on air quality in the European Union not only force member states to monitor and report on their air quality but also to actively inform the public on the status of the ambient air quality. The Aarhus convention, ratified by the EU in 2005, further enforces the concept that citizens have the right to be informed on the environmental conditions they live and work in. Over the past years a good number of cities and countries have started to display monitored or modeled air quality data on the internet. For most of the monitoring organizations, the internet is the easiest way to meet the dissemination of information requirements of the European (and/or national) legislation. The fact that so much air quality information is available on the internet makes it tempting to compare different cities in different countries. This proves particularly difficult. Apart from technical websites such as the European Environmental Agency's ozone website (http://ozone.eionet.eu.int) and Airbase (http://airbase.eionet.eu.int) there are no possibilities to compare cities/countries side by side. Even if one surfs from one site to the other, comparison is not easy: air quality is presented in different ways using different interpretation criteria and a different typology of stations, which is usually not clearly explained.

The most widespread way to interpret air quality on websites is the use of an index ranging from good to bad to make the detailed measurements in micrograms more understandable for the general public. A review of existing websites and the associated air quality indices shows that the way air quality is interpreted differs considerably across the world. More surprisingly, even amongst the EU member states who share common legislation, the indices used do vary. There are a number of reasons to explain these differences. Some of them are historical, others conceptual: the index existed before the EU regulations came into force or the index is based on health and exposure criteria, e.g. the UK index (DEFRA, www.airquality.co.uk/archive/standards.php#band). The fact that air quality problems (sources, meteorological conditions, etc.) tend to differ is also one of the reasons. The indices tend to be calibrated to the local situation, for example, to make sure that there is some variation in the index from day to day (to entice repeated visits) and that the typical range of pollutant conditions occurring locally is being covered.

To facilitate the international comparison of near real time air quality the CITEAIR project (co-funded by first by INTERREG IIIC then INTERREG IVc, http://citeair.rec.org) has developed a common operational website www.airqualitynow.eu where cities can display their air quality information side by side. The project aims at making air quality comparable across Europe and www.airqualitynow.eu is already open for any city to join. The common website relies on common air quality indices (hourly, daily and annual). The comparison possibilities offered to the public are on an hourly, daily and yearly basis and the indices were developed keeping in mind that the general public is the end user. This paper presents the CITEAIR common website and its set of indices. It explains how they have been elaborated through collaboration between the partner cities (Leicester, Paris, Prague, Rome and Rotterdam) and is now open to any city willing to participate.

1.2 Difficulties in comparing the European cities air quality

CITEAIR provides a common index and makes a difference between background and traffic stations. The potential of having a common index on the same website will be apparent from the following example in which we try to compare air quality at a given day in four cities. The indices are described in table 1. Three out of four cities have an index, two indices range from 1 to 10, the other from 1 to 200. Two cities have 10 classes, one has 5, one has 4. Two describe air quality in terms of good and bad, one in terms of health and the fourth in terms of pollution levels. The class boundaries are very different. In addition, whereas most websites have a page explaining how the index is calculated, other methodological aspects are generally not explained. Does the index represent measurements at background stations, traffic stations, a mixture? If someone would like to compare these four cities at a given moment he or she would not only have to visit four websites but also be faced with four completely different presentations, qualifications and languages. Apart form the fact that the bands differ from one country/city/area to the other, the data behind the index also differ. This is hard for a specialist and almost impossible for the general public. The CITEAIR initiative comes from this statement.

Table 1: Indices used on the internet in Paris, Leicester, Rome and Rotterdam (2005)

ATMO Paris	ozone-1h	PM10-24h	NO2-1h	index	UK	ozone-8h	PM10-24h	NO2-1h	index
Very good	29	9	29	1	low	32	21	95	1
	54	19	54	2		66	42	190	2
good	79	29	84	3		99	64	286	3
	104	39	109	4	moderate	126	74	381	4
average	129	49	134	5		152	86	477	5
mediocre	149	64	164	6		179	96	572	6
	179	79	199	7	high	239	107	635	7
poor	209	99	274	8		299	118	700	8
	239	124	399	9		359	129	763	9
Very poor	>=240	>=125	>=400	10	very high	>=360	>=130	>=764	10

Rome	ozone-1h	PM10-24h	NO2-1h	index	Rotterdam *	ozone-1h	PM10-24h	NO2-1h	index
good	90	100	100	50	good		20	100	-
moderate	135	150	150	75	moderate	180	40	200	-
mediocre	180	200	200	100	bad	240	60	400	-
unhealthy	360	400	400	200	very bad	>240	>60	>400	-
very unhealthy	> 360	> 400	> 400	>200					

* Ozone classification: national smog pages- Other classes :local traffic website.

1.3 Comparing European cities thanks to a common website

The CITEAIR common website www.airqualitynow.eu has been launched in March 2006. It is meant to be an interesting complement to the cities own websites. The website and the corresponding indices are not launched to replace existing websites or indices. Their value added is to provide for the first time a European and comparable picture of the air quality at first glance, up to date and understandable by anybody for:

- three time scales (on a hourly, daily and an annual basis),
- and two types of public exposure to air pollution (background and traffic conditions).

In addition, for those cities that are not yet on the internet, and/or do not currently use an index, "air quality now" and its indices could be their primary platform to easily provide information to the public and the local authorities. A space is also offered to cities for presenting themselves according to a common template, providing background information on their specific air pollution situation and associated reduction measures.

The website is operational with currently almost 40 cities feeding their data (see fig. 1). The process to join has been made easy. Based on the data sent by the cities through an agreed ftp format, the indices calculations are automatically done inside www.airqualitynow.eu. The full procedure is detailed on the "join us" page. However, not every city has its own monitoring network or both traffic and background stations, and not everyone is able to deliver data in near real time. If cities want to participate in only one of the indices, can only deliver data on a daily basis, or even only present year average data, they can still join the website. Different sections of the website provide a platform to compare different data (year average, daily, hourly). Participation is therefore not limited to those with their own automated monitoring network. Other cities are invited to join and upload their data as well.

Www.airqualitynow.eu does not aim to replace more targeted existing local information. This would be an unrealistic ambition as in many cities the public has got used to the local, tailor-made index. The proposed common indices are, by the nature of the fact that they are common to a wide area, a non-specific compromise. CITEAIR envisages that there is room for two sources of air quality information on the internet: a local website, in the national language with a dedicated presentation (often using a well established and known local index); and a common website aimed at comparing - in near real time - the air quality in your own city to the air quality in other European cities.

Figure 1: Home page of the CITEAIR common web site www.airqualitynow.eu presenting the comparison of air quality in European cities through common indices (background situation for the 14/01/09 at 01:00 pm).

2. INTERNATIONAL COMPARISON OF CITIES THANKS TO COMMON AIR QUALITY INDICES

2.1 Review of air quality indices

There are a number of different ways to interpret air quality in near real time. The most common way to do so is the use of an index, generally based on a number of sub-indices for individual pollutants. Air quality indices aim to translate the chemical characteristics of a quite complex mixture of pollutants in the air into one single figure. From a scientific point of view this is obviously a gross generalisation but for communication purposes it is considered an essential simplification. An index is also always a compromise between several objectives and potentially occurring situations. The trade-offs in developing the common indices were made having in mind that they should be applicable over a wide range of conditions and interesting to the public.

There is a wealth of indices and countries that share the same legislation, or sometimes even areas/cities within the same country, have different indices as shown by Garcia et al. [2002] and van den Elshout et al. [2007]. Some of the differences can be explained by the local differences in the nature of the air quality problems. Some other differences are due to fundamentally different approaches. For example, the UK and US-EPA indices (http://cfpub.epa.gov/airnow/index.cfm?action=aqibroch.index) are strongly related to perceivable health effects. The bands in the index are explained in health terms. This implies that the index covers a very wide range of concentrations and that the observed concentrations are very often in the "good" or "moderate" end of the scale. Air quality in Europe, fortunately, is rarely poor enough to cause acute health effects so any index based on health impacts tends to trail at the lower end of the scale for most of the time.

Other indices take a different approach. For example the ATMO index, based on a national regulation concerning all French cities larger than 100 000 inhabitants (ADEME, www.buldair.org) has bands that are somehow linked to values that are also used in the current EU directives. The alert thresholds in the directives tend to define the higher end of the scale. In these cases the top end of the index scale ends somewhere in the middle of the health effect based scales. For example the worst end (very poor) of the NO_2-index in France corresponds to 400 µg/m3. In the UK this is in the lower end of the "moderate" band and in the US it is even considered too low to calculate an index value.

Communication-wise the health-based indices have both a clear advantage and a disadvantage. The advantage is that the index value displayed at the website is easy to interpret: it does or does not cause health effects. The disadvantage is that the index is almost always indicating that air quality is good and pollution is low whereas the limit values for long-term exposure are often exceeded. This leads to an apparent paradox: a citizen regularly checking the local air quality website will always get the message that the air quality is good whereas at the end of the year local government puts out a report that he or she is living in a hotspot area for which an action plan is required. This is the paradox between short- and long-term air quality criteria. The criteria for short-term exposure are often met except for episodes, like for example in the summer of 2003. The criteria for long-term exposure are often not met in Europe's urban areas. The ATMO-type of indices provides some differentiation at the lower end of the scale to assure that the air quality is not always "good". However in this case it is very difficult to attach some kind of health interpretation to the index.

The long-term vs. short-term paradox typically occurs on the internet. In an annual report the focus is on long-term air pollution and in the teletext pages dedicated to smog warnings the focus and interpretation is based on short-term health effects. However, internet presentations often serve multiple roles: informing the public, but also making the public aware of air quality issues. In this case the paradox is difficult to resolve: highly variable hourly (or daily) data is being presented to assure an attractive and frequently changing presentation that encourages repeated visits. On the other hand, the most challenging limit values appear to be the criteria for the year average so interpreting commonly occurring hourly values in terms of good or bad is fairly arbitrarily. They are not bad from the short-term exposure point of view but might be bad from the long-term exposure point of view. This explains the need for an annual index as well.

2.2 A common daily and hourly index (CAQI)

As already explained, comparing air quality in different cities is a tricky issue. Is the air quality being determined in the same way (this mainly applies to particulate matter) and at comparable locations? This is not an issue that the CITEAIR project and its indices can solve. Its common website will take for granted whatever a city supplies as input in either category. As a first step to improve comparability, the indices will be reported both for roadside and city background locations. This is considered an important improvement over city averages: some monitoring networks are designed to monitor or spot areas of poor air quality (with possibly a high number of roadside stations) whereas others are aimed at providing an average city picture.

The Common Air Quality Index (CAQI) is used both for a daily index and for an hourly index. In the website the daily index will be shown for the past day (D-1). For the current day, the hourly index will be available, to be updated every hour. A daily index for today would need forecasting or 'nowcasting' a facility that is not available in each city with a monitoring network, hence the option of an hourly index. The hourly index is also a reasonably dynamic parameter, enticing repeated visits to a website.

The CAQI is calculated according the grid in table 2, by linear interpolation between the class borders. The final index is the highest value of the sub-indices for each component. As can be seen there are two CAQI-s: one for traffic monitoring sites and

Table 2: Pollutants and calculation grid for the CAQI

Index class	Grid	Traffic				City Background					
		Mandatory pollutant			Auxiliary pollutant	Mandatory pollutant				Auxiliary pollutant	
		NO_2	PM_{10} 1-hour	PM_{10} 24-hours	CO	NO_2	PM_{10} 1-hour	PM_{10} 24-hours	O_3	CO	SO_2
Very low	0	0	0	0	0	0	0	0	0	0	0
	25	50	25	12	5000	50	25	12	60	5000	50
Low	26	51	26	13	5001	51	26	13	61	5001	51
	50	100	50	25	7500	100	50	25	120	7500	100
Medium	51	101	51	26	7501	101	51	26	121	7501	101
	75	200	90	50	10000	200	90	50	180	10000	300
High	76	201	91	51	10001	201	91	51	181	10001	301
	100	400	180	100	20000	400	180	100	240	20000	500
Very High*	> 100	> 400	>180	>100	>20000	> 400	>180	>100	>240	>20000	>500
NO_2, O_3, SO_2:	hourly value / maximum hourly value in $\mu g/m^3$										
CO	8 hours moving average / maximum 8 hours moving average in $\mu g/m^3$										
PM_{10}	hourly value / daily value in $\mu g/m^3$										

* An index value above 100 is not calculated but reported as "> 100"

one for city background sites.

The traffic index comprises NO_2 and PM_{10}, with CO as an auxiliary component. The background index obligatory comprises NO_2, PM_{10} and O_3, with CO and SO_2 as auxiliary components. In most cities the auxiliary components will rarely determine the index (that is why they are auxiliary) but in a city with industrial pollution or a seaport SO_2 might occasionally play a role. Benzene is considered a long-term exposure issue. The number of cities with online monitoring benzene is limited and it is therefore not included in the short-term indices.

The choice of the classes in the CAQI is inspired by the EU legislation and based on a compromise between the participating cities. The dividing line between medium and high is often linked mainly to the values mentioned in the directives: alert thresholds (SO_2, NO_2, O_3) or air quality objectives when available on a daily basis (CO and PM_{10}). Class borders were regularly spaced for the main components. PM10 is an exception. To avoid that the CAQI is completely dominated by PM_{10} the value of 50 µg/m^3 as a daily average was positioned as the bordering line between low and medium. For the setting of the CO and SO_2 borders additional inspiration was sought from Cairncross and John and the DAPPS index [2004] which aims to define the component sub-indices based on the relative risks attributed to each component.

The CAQI resembles the French ATMO index discussed above and it differs substantially from for example the UK and US-EPA indices. It therefore shares the drawbacks of the ATMO :no clear link with health effects, fairly arbitrarily quality interpretation of hourly values. But it also shares its advantage: frequently changing index values that capture the hour-by-hour changes and make a website dynamic. The latter was of overriding importance as raising awareness is a key objective of the common website. De Leeuw and Mol [2005] compared the CAQI to a number of other indices.

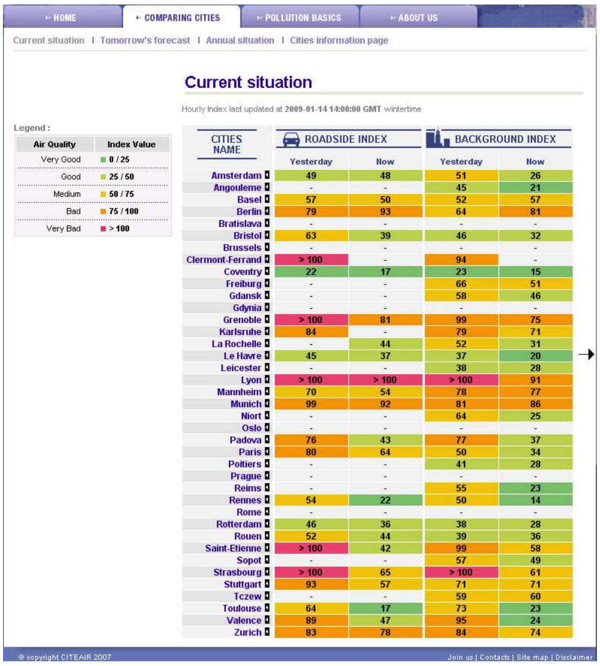


Figure 2: CAQI results for the cities part of CITEAIR and details of the CAQI results for the agglomeration of Paris on www.airqualitynow.eu.

Whereas the calculation grid for hourly and daily values is the same for most components, PM_{10} poses a particular problem. Consistency had to be found between the hourly and the daily index. Many networks only report 24-hour (moving) average data. This different averaging time implies that their concentration readings are always lower than those for the networks reporting true hourly values. In order to solve this problem, a selection of 52 urban and suburban monitoring stations from Airbase for the period 2001-2004 has been used to calculate the ratio between daily maximum hourly concentration and daily average concentration. This average ratio appears to be 0.55. Based on a wide selection of stations, it is used to link the hourly and daily index grid.

Figure 2 presents how the CAQI is displayed on www.airqualitynow.eu for the almost 40 current participating cities and the possible comparisons for a particular day and hour enabled by the index for two time scales and two types of exposure. As an example, the details of the CAQI calculation for Paris through its sub-indices are also provided.

2.3 A common year average index (YACAQI)

Year average indices are not very common in air quality reporting but they are nevertheless a useful indicator for non-experts, facilitating the comparison of cities at a glance. Comparing cities by their individual pollutant levels is difficult as one city might be better on one pollutant and worse on the other. In addition, some cities might monitor other pollutants than others. Even comparing progress in a single city from one year to the other is difficult as progress might be made for one pollutant whereas in another field things might have deteriorated. A year average index is a huge simplification but it does provide an easy way to make some kind of relative assessment on the position of one city to the other or for one city from year to year.

In CITEAIR the way of making such a medium term index is the "distance to target" principle. One advantage of the distance to target principle is that each parameter considered contributes to the index (unlike the principle where the worst parameter determines the index). A distance to target indicator calculates, for each pollutant, a ratio of how far the actual measurement is away from the target value, for example a limit value. The overall index/indicator is the average of the sub-indices. A distance to target index is based on policy targets or limit values. The limit and target values have important implications both for environmental policy makers and for the public. Besides, they do have a link to health risks: in Europe they are most of the time related to the recommendations of the World Health Organisation [2005]. The distance to target the way of making an index is the year average index presented in this paper and used on www.airqualitynow.eu.

Table 3: Calculation basis for the year average index

Pollutant	Target value / limit value	Calculation
NO_2	Year average is 40 µg/m^3	Year average / 40
PM_{10}	Year average is 40 µg/m^3	Year average / 40
	Max. number of daily averages above 50 µg/m^3 35 days ≈ year average of 31 µg/m^3	Year average / 31
Ozone	25 days with an 8-hour average value >= 120 µg/m^3	# days with 8-hour average >=120 / 25
SO_2	Year average is 20 µg/m^3	Year average / 20
Benzene	Year average is 5 µg/m^3	Year average / 5
CO	-	Not calculated

Like the hourly and daily index, the Year Average Common Air Quality Index (YACAQI) is calculated for traffic and city background sites. Preferably a city's data for each index is based on the average of a number of sites, however it is up to each city what they want to contribute (monitored or modeled data). In most cases, the data provided to the website will be based on the situation at one or more monitoring sites.

The calculation of the sub-indices is detailed in table 3. Sub-indices are calculated for each pollutant by dividing the actual year average by the EU limit or target value. The overall city index is the average of the sub-indices for NO_2, PM_{10} (both year average and daily averages) and ozone for the city background index. For the traffic year average index the averages of the sub-indices for NO_2 and PM_{10} (both year average and

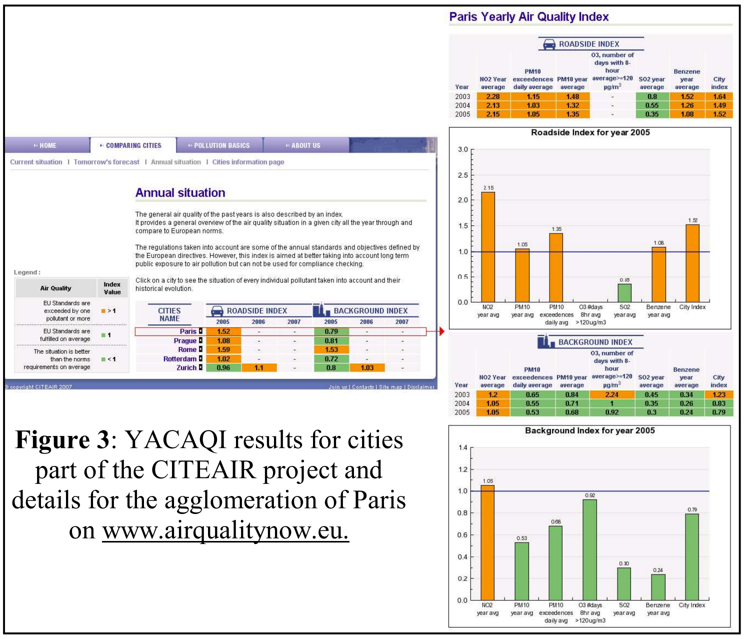


Figure 3: YACAQI results for cities part of the CITEAIR project and details for the agglomeration of Paris on www.airqualitynow.eu.

daily averages) are being used. The other pollutants, if data are available, are used in the presentation of the YACAQI but do not enter the calculation of the city average index. They are treated as additional pollutants like in the hourly and daily indices. The main reason is that not every city is monitoring the full range of pollutants. Furthermore for SO_2 we expect that the situation in different kinds of cities is very far apart, being no problem in most cities and a concern in others. Figure 3 presents how the YACAQI is displayed on www.airqualitynow.eu for a selection of cities, for two types of exposure, over a number of years, and the details for Paris. This type of presentation provides valuable additional information when comparing two cities or the same city over two years. At a glance it becomes evident what the main problems are and where progress for the situation is satisfactory.

3. CONCLUSION

Full details on the elaboration of CAQI and YACAQI, as well as sample application, are available on www.airqualitynow.eu. Maps of forecasted concentrations of the pollutants of main concern have already been added and will be complemented by forecasted air quality indices during the life time of CITEAIR II. Further developments of these common tools will concern: participation of new cities and new media partnerships to display the indices, integration of $PM_{2.5}$ in the indices according to the CAFE directive, as well as translations in several EU languages.

Cities are engaged in communication with the public, not only because of legal obligations but also to raise awareness. This implies that air quality issues have to be presented in an attractive and educational way. The possibility to compare your own local air quality to a number of other European cities could be an asset in this respect. The purpose of common indices and website is not to replace more detailed local information nor to check EU regulation compliance but to complement it. The value added is to provide, for the first time, a European and comparable picture of the air quality, near real-time and understandable by anybody. In addition, the provision of separate indices for two types of environmental conditions and three time scales is a methodological innovation. It could also be an alternative to raise public awareness for cities which do not already operate a website.

ACKNOWLEDGEMENTS

CITEAIR II consists of 11 partners (14 partners for the previous project) from 8 European countries. This work is the result of contributions from all of them and supports from INTERREG (IIIc than IVc), the users community as well as the EEA.

REFERENCES

Air quality guidelines for PM, O3, NO2 and SO2: summary of risk assessment, World Health Organisation (WHO), http://www.who.int/phe/air/aqg2006execsum.pdf, 2005

Cairncross, E. K. and John, J,. Dynamic Air Pollution Prediction System (DAPPS). *Communicating air pollution exposure: a novel air pollution index system based on the relative risk of mortality associated with exposure to the common urban air pollutants*. IUAPPA 13th Annual World Clean Air and Environmental Protection Congress and Exhibition. London, 2004.

Garcia, J., Colossio, J. and Jamet, P. *Air-quality indices. Elaboration, uses and international comparisons*. Les presses de l'Ecole des Mines, Paris, 2002.

Leeuw, de Frank, and Mol, Wim.. *Air quality and air quality indices: a world apart*? ETC/ACC Technical paper 2005/5.

Van den Elshout, Sef and Léger, Karine. Comparing Urban Air Quality Across Borders. A review of existing air quality indices and the proposal of a common alternative. http://citeair.rec.org/downloads/Products/ComparingUrbanAirQualityAcrossBorders.pdf 2007

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Integration of Earth Observation Data and Domain Knowledge Using Topic Maps

Rani Pinchuk[a], **Bernard Valentin**[a], **Samjwal Ratna Bajracharya**[b], **Rajan Man Bajracharya**[b]

[a]*Space Applications Services, Leuvensesteenweg 325, B-1932 Zaventem, Belgium*
({rani.pinchuk, bernard.valentin}@spaceapplications.com)

[b]*ICIMOD, GPO Box 3226, Kathmandu, Nepal*
({sabajracharya, rbajracharya}@icimod.org)

Abstract: A novel approach for integrating of Earth Observation (EO) data with domain knowledge using Topic Maps is presented. The domain of the research is glacier, glacial lake and the glacial lake outburst flood (GLOF) phenomena. Much information is available over thousands of glaciers and lakes in Nepal as well as relevant EO data of the region. However, because the data is available from different sources, it is not simple to access it. In the project, two Topic Maps ontologies are designed: one for describing the domain data and the other for describing the EO data. The merging of the two ontologies provides one semantically rich access point for all the domain data and relevant EO data. After introducing the approach, the paper describes the domain of research, and then introduces the Topic Maps technology. Following that, the ontologies designed in the project and the software architecture are described. Next, several other environmental domains where the approach described may be useful are listed, and finally conclusions and future work are presented.

Keywords: Environmental knowledge; Earth Observation; Topic Maps; Ontology; Glacial Lake Outburst Flood; Global Warming.

1. INTRODUCTION

Earth Observation data is valuable for understanding different environment related issues such as floods, forest fires, pollution, or earthquakes. However, in order to allocate a specific EO product (e.g. satellite image) which relates to the problem at hand, different domain knowledge data stores have to be accessed. These data stores can be multidisciplinary or interdisciplinary and are often fragmented across many organizations and projects. Accessing different data stores requires the knowledge of different user interfaces and database designs and leads to inefficiency.

For example, a researcher studying Glacial Lake Outburst Floods (GLOFs) has first to select the sort of glacial lakes (moraine dammed, ice dammed, supraglacial, erosion, etc.) that are of interest. Other factors related to glacial lakes should be selected as well (drainage condition, orientation, elevation, etc.) Then the location of each such lake should be identified, and finally the relevant EO products should be allocated. When doing, for example, a comparison study, this process can be tedious. This inefficiency is emphasised even more when considering the impact of the weather fluctuations on the glaciers. A researcher that wants to understand the processes involved has also to navigate through the weather data in yet another database.

It is clear that integrating databases of EO products with databases containing domain knowledge provides a great advantage for domain experts and decision makers. The user has one access point to the information and can use the domain terminologies.

Topic Maps technology is an ISO standard (ISO/IEC 13250:2003) for the representation and interchange of knowledge. It is especially useful in the above scenario for several reasons:
1) Topic maps are built from topics and associations between them. A topic represents a subject and can contain occurrences that relate to this subject (a reference to text about the subject, a reference to an image describing the subject, etc.) In addition, topics can be associated to each other. These associations provide a structure that corresponds to the way we conceptually grasp knowledge. Such a structure provides a way to navigate between the topics in the topic map. It also lets users ask questions about the knowledge within the topic map, for example: where a certain glacier is located, or what is common to a group of glaciers. This gives meaning to the data that is kept in the topic map, and therefore Topic Maps technology provides a semantic layer over data and facilitates its accessibility.
2) As mentioned in (1) above, each topic can contain occurrences that are data or references to resources that are related to this topic. These references can point to any multimedia resource such as voice, image, video, etc. This makes Topic Maps an excellent choice for creating an advanced meta-data layer over multimedia resources.
3) The Topic Maps standard provides the ability to merge topic maps. This means that when a topic map that describes the EO products (where, for example, each product is represented by a different topic, and those topics are associated to a mission, to a period, etc.), and a topic map that describes glaciers and glacial lakes are given, these two topic maps can be merged in order to achieve an integration of the resources.

In the SATOPI project, this approach is demonstrated in a use case targeting the GLOF phenomena. ICIMOD (International Centre for Integrated Mountain Development) has collected data over 3252 glaciers and 2323 glacial lakes in Nepal. This data includes different characteristics of the lakes such as physical condition, spatial data, orientation or classification. In addition, EO data from different sensors is available over the region. The project provides one access point to the available data, which allows interaction with the data using domain specific terminologies. This is done by authoring two topic maps: A first Topic Map Ontology is designed to include the above described data about glaciers and glacial lakes and a topic map is created by populating this ontology with the data from the existing resources or by connecting the resources to the topic map (shallow copy). In the same manner a second topic map is created which represents the EO data. The two topic maps are merged and a user interface is provided in order to query the overall data using domain specific terminologies.

2. THE DOMAIN OF RESEARCH – GLACIER, GLACIAL LAKE AND GLOF

ICIMOD initiated the development of the database of glaciers and glacial lakes in different basins in the region in a systematic manner with unified inventory approach. Due to harsh climatic conditions and difficult terrain and accessibility in terms of road infrastructure in the areas of the Hindu Kush-Himalayan (HKH) region, remote sensing (RS) plays a vital role in carrying out the development of such an inventory.

The inventory data gives information on glaciers, glacial lakes, areas where GLOF events have occurred and lakes that could pose a potential threat of a GLOF in the near future. This documents the basis for the development of a monitoring and early warning system and for the planning and prioritization of disaster mitigation efforts that could save many lives and properties situated downstream, as well as guide infrastructure planning. In addition, this database could be useful for addressing many other environmental issues such as environmental hazard, flood forecasting, etc.

A glacial lake outburst flood or GLOF – a form of sudden flash flood – often has catastrophic effects downstream. GLOF events pose severe geo-morphological hazards and their floodwaters can wreak havoc on all human structures located on their path. GLOF events have resulted in many deaths, as well as the destruction of houses, bridges, fields, forests and roads. Unrecoverable damage to settlements and farmland can take place at great distances from the outburst source. In most of the events the livelihood is greatly disturbed for long periods.

A number of GLOF events have been reported in the region in the last few decades and an increase in their frequency is anticipated in coming years as many glacial lakes are known to be formed in the Himalaya in the last half century due to global warming. The immediate impact of the accelerating global warming was the rapid glacier melting. Melting glaciers (high frozen reservoirs) releases their water at the top of their watersheds that wind their way through thousands of kilometers of grazing, agricultural, and forest lands and are used as renewable sources of irrigation, drinking water, energy, and industry. The glaciers are particularly vulnerable to climate change as rapid accumulation of water in glacial lakes can lead to a sudden breaching of the unstable moraine dams.

However, the information collected is kept across many different platforms in a form of database, metadata, publications, articles, scientific findings, websites, literature and bibliography etc. The benefits of having a single access point to all these resources are obvious. It is also important to understand that information is the combination of knowledge and observations. To be scientifically useful, Earth Observation data must be combined with knowledge in order to improve our understanding of the Earth system.

3. THE TECHNOLOGY USED – TOPIC MAPS

As mentioned above, Topic Maps is a standard (ISO/IEC 13250:2003) for knowledge representation and information integration. It provides the ability to store, together with the data, complex meta-data that represents the meaning of the data stored.

In Topic Maps, each subject in the knowledge domain is represented by a topic. In Figure 1, "Imja Tsho" is a topic representing the Glacial Lake Imja Tsho. Each topic may have a type. Here, the topic "Imja Tsho" is typed by the topic "Lake", meaning that Imja Tsho is a lake.
Topics can be associated with other topics. The topic "Imja Tsho" is associated to the topic "Koshi River Basin" and the association between them is "located-in". The fact that topics are associated with each others provides a structure that corresponds to the way we conceptually grasp knowledge. Such a structure provides a way to navigate between the topics in the topic map. It also lets users ask questions about the knowledge within the topic map (for example: "Where is "Imja Tsho" located?"). This gives meaning to the data that is kept in the topic map.

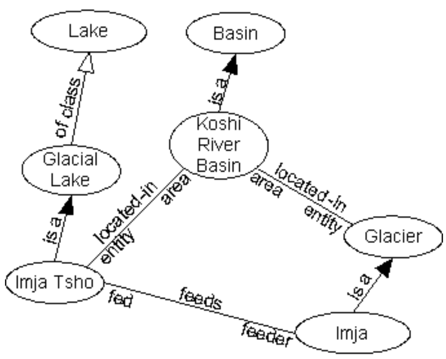


Figure 1. Topic Maps example

Note that associations are built from players, roles and association types. In the example given above, "Imja Tsho" plays the role "fed" in the association of type "feeds" and "Imja" plays the role "feeder" in the same association.

Each topic has one or more names (e.g. "Imja Tsho"). In addition it can hold extra information called occurrences. These are pieces of information about the topic, similar to the paragraphs in a book that are referenced by a book's index. Occurrences can be simple textual information or references to any other media (such as web pages, audio, images, video, etc.) Like associations, occurrences are also typed. Therefore the topic "Imja Tsho" might have for example an occurrence of type “mean length” (“410 meters”) or of type “Article” (where the actual URL of the article is kept as the occurrence value).

Different Topic Maps constructs – names, occurrences and associations – might be scoped by other topics. The scoping topic describes when the construct is in scope (or is out of scope). For example, we can provide more than one occurrence of type “mean length” for a lake. Each such occurrence is correct in a certain period. We can define topics which represent those periods and use them to scope the different occurrences.

All the types in a topic map – the topic types, the occurrence types, the association types and the role types – are defined as topics. These topics provide the conceptual skeleton of the topic map. These topics together with the scoping topics are referred to, in the Topic Maps community, as the Topic Maps Ontology. Ontologies are very useful when authoring topic maps as they help to identify the borders of the domain of knowledge that the topic map represents.

4. THE SATOPI IMPLEMENTATION

4.1 Description of the Ontology

The Topic Maps ontology is designed by the domain experts with the support of the project analyst. This designing process has proved to be very useful in understanding the domain in question, the data available and the user requirements. During the SATOPI project definition phase, the domain experts provided the project analyst with a description of their domain of activity: the different types of entities, the relations that exist between those entities and the kind of information they have in hand about those entities.

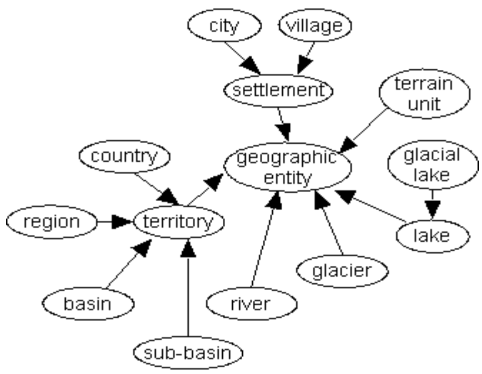


Figure 2. Top level types

In the first stage, the different top-level topic types were identified. At this stage it was also understood that two different topic maps will be created – for the domain knowledge data and for the available EO data. Figure 2 describes the top level types for the first topic map.

The next step is to describe in details the different topics of the different types. For example, in Figure 3, the different occurrences and relationships for topics representing glaciers are presented.

In the same manner the structure for all the main topic types is designed.

Another issue that has to be solved at this stage is the policy in which the different topics in the topic map are uniquely identified. Each topic in the topic map can have one or more unique identifiers. These identifiers allow for successful merging between topic

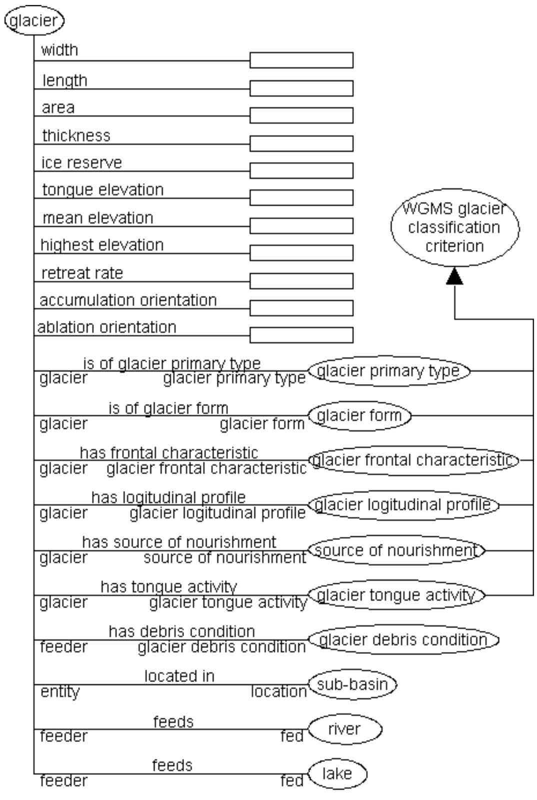


Figure 3. Glacier topic type

maps. Usually during the design of the ontology, a rule for assigning these identifiers is set per topic type. For example, in the SATOPI project, all the identifiers of glaciers are built from ids given by ICIMOD, so the identifier of the Imja glacier is http://satopi.icimod.org/d/glacier/G086986E27651N/.

4.2 The Architecture

The backend of the application is built around a Topic Maps engine called TopiEngine. A Topic Maps engine is an engine which provides an API for accessing and manipulating topic maps. Usually the engine also provides an implementation of the Topic Maps Query Language (TMQL). However, the data in the topic maps of projects such as SATOPI is spread across different data stores. Therefore the engine has been extended by data store connectors. Each such connector is generated automatically from a mapping definition file. This file defines how the data store can be accessed and how data in it is mapped to the ontology of certain topic map residing in the Topic Maps engine. This way the engine provides a Topic Maps view of data collected from different sources. This whole structure is wrapped as an Apache module which means that the frontend can access the data without taking into account the complexity of the location of the different parts of the data available.

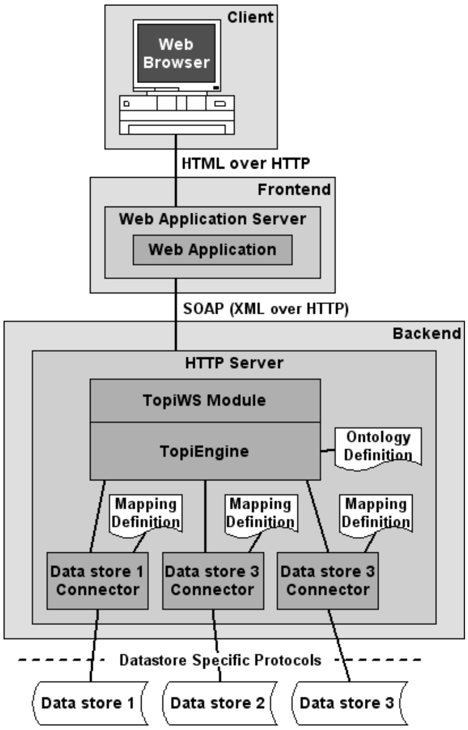


Figure 4. SATOPI architecture

Moreover, different parts of the data can be provided as different topic maps. For example, the EO data can be provided in a topic map which contains only the relevant EO data and locations, but does not include glaciers and GLOFs related information. This information can be accessed through a second topic map. The ability to merge those two topic maps provides a second level of federation of the data. In other future applications, one of those topic maps can be re-used. For example, the same EO data can be used in different research in the same geographic area. The glaciers and GLOFs related topic map can be merged in another application related to the water supply in the region of the Himalaya Mountain.

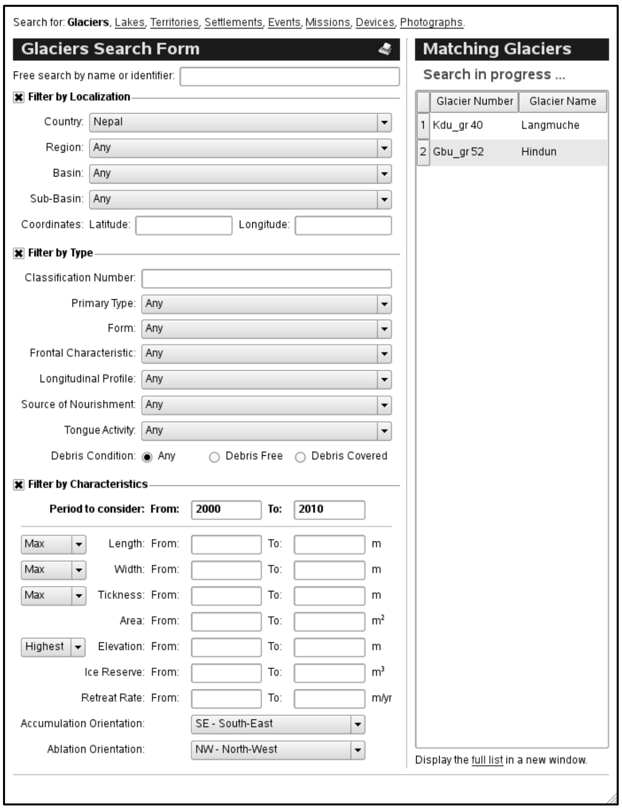


Figure 5. Glaciers Search Page

Because data provided in topic maps is organized in a very meaningful way, it is much easier to provide a rich user interface. In implementing the user interface in the SATOPI project, the Skin design pattern has been implemented. Different templates (skins) where created for the main topic types in the application. They were chosen to be Glacier, Lake, Territory, Settlement, Event, Mission, Device and Photograph.

For each of these types, a dedicated search page is provided. Figure 5 shows the search page that allows the users to find glaciers using various criteria. Figure 6 shows how a specific topic is presented to the user. This page about the Imja Tsho glacial lake provides the user with full information about the subject presented.

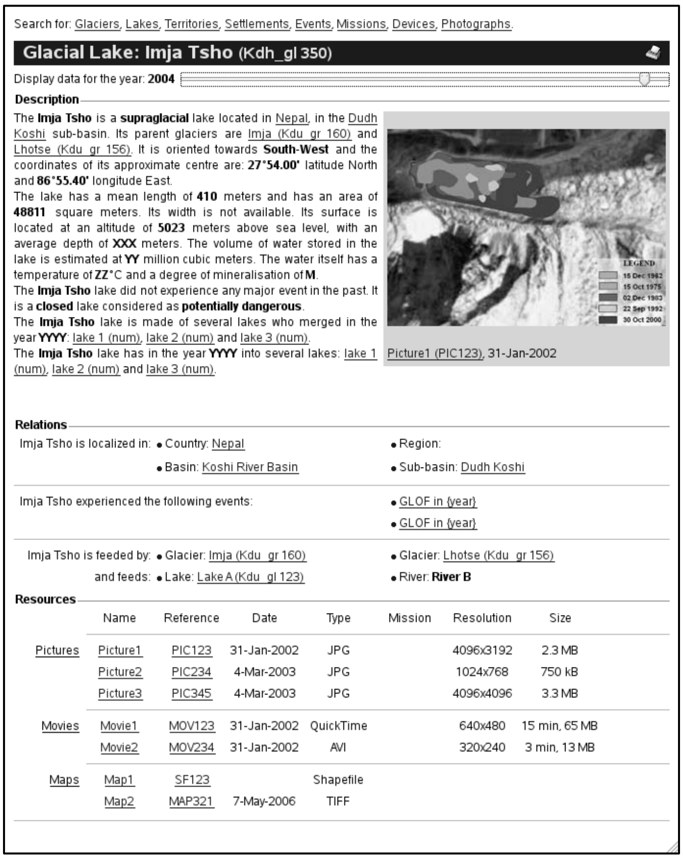


Figure 6. Glacial Lake description page

Both pages demonstrate the strength of using Topic Maps and templates. Because the data is organized according to meaningful types, different subjects are presented in different ways which match the special considerations for the subjects of those types. This is translated to the use of a specific template for a specific type. The templates themselves are easy to define because the data can be addressed by the template author in a meaningful way. For example, the location of a glacier lake is associated to the topic representing the lake with the association "located-in". This facilitates the authoring of the templates and allows to provide a sophisticated and rich user interface.

5. OTHER DOMAINS WHERE SATOPI TECHNOLOGY MAY BE USED

The expected global changes and the associated impacts will make Earth Observation-derived data increasingly vital for supporting sustainable development. The EO application is diverse ranging from cryosphere monitoring, land cover study, hazard and vulnerability mapping, aerosol depth study, satellite rain fall estimation, weather forecasting, etc. and increasingly recognized as an indispensible tool for monitoring, assessment and planning of natural resources and the environment.

More specifically, SATOPI-like solutions where EO data is integrated with domain knowledge can be seen on three different areas as mentioned below.

5.1 Cryosphere

Cryosphere collectively describes the portions of the Earth's surface where water is in a solid form and includes sea ice, lake ice, river ice, snow cover, glaciers, ice caps and ice sheets, and frozen ground (which includes permafrost). The cryosphere is an integral part of the global climate system with important linkages and feedbacks generated through its influence on surface energy and moisture fluxes, clouds, precipitation, hydrology, and atmospheric and oceanic circulation. It is due to these feedback processes and amount of freshwater content that makes cryosphere a vital environmental component in global climate modelling and eco-system services.

For further impact analysis imposed by glaciers, glacial lakes and GLOFs, a need of real time or near real time data is very important and will be beneficial for the community and to carry out environmental studies. Tremendous amount of data will thus be generated as daily to weekly products of snow cover map, snow depth map, snow water equivalent map. It gives information on the snow cover extent, depth and the amount of water that would result when completely melted, thus making the product extremely important for research themes such as snow cover change, river flow regime, water resource assessment, and flood management. However, methodologies to derive these products need to be adapted to mountain specific conditions and need to be validated with in-situ measurements. Finally, the data needs to be disseminated to the user community through an integrated approach by combining data, information and knowledge together.

Therefore, there is a need to develop a system that serves as a gateway to the host of information required for research. Here, the SATOPI project can be further extended by integrating cryosphere related products and information. This is especially relevant to an organization like ICIMOD, striving to be an enabling, learning and knowledge centre.

5.2 Land Cover

Land cover is an important aspect that constitutes the basic core information on natural resources at a national and regional scale. Land cover is highly relevant for data scarce region such as the Hindu Kush-Himalayan region with high degree of inaccessibility. EO applications can support periodic assessment of natural resources to study the land cover and land use dynamics at multiple scales. Data harmonization among the Regional Member Countries (RMCs) and standardized methodology for land cover classification system and technology transfer to the RMCs will be the key elements. Linking with global land cover initiative by FAO and forest cover assessment will provide opportunity for ICIMOD as a regional platform to contribute to the global efforts. Regional land cover mapping, vegetation index, carbon sequestration are key application and products.

Integrating these products together with domain knowledge as a SATOPI-like platform will provide value added service by enabling access to its resources and knowledge for a wider audience.

5.3 Atmosphere

Atmospheric aerosols interact with sunlight and affect the global radiation balance. It is known that atmospheric aerosols affect climate not only directly by scattering and absorbing visible and infrared energy, but also indirectly by modifying the properties and the lifetime of clouds. Especially the Asian continent is one of the most important sources of aerosols.

Since the lifetime of atmospheric aerosols is from few days to few months, their spatio-temporal distribution is highly variable. Satellite remote sensing techniques have the advantage of being capable of measuring optical and physical aerosol parameters over a large area. These techniques also provide the spatial and spectral resolution necessary to monitor the highly variable aerosol pattern. Various algorithms are used to estimate satellite-retrieved aerosol optical depth (AOD) over both ocean and land. Satellite retrieved AOD provides spatial distribution of aerosols to estimate the impacts of atmospheric aerosols on air quality and radiation. They are capable of monitoring severe air pollution episodes such as biomass burning and Asian dust. Especially in the Hindu Kush-Himalayan region where the terrain is difficult, this technique would be useful to make the regular monitoring and estimate its optical property to study the transport and impact of the anthropogenic aerosols.

As the aerosols play a significant role in global climate change. Measurements by satellites provide important insights to study the impact of trans-boundary air pollution from local to global level. ICIMOD houses an Atmospheric Brown Cloud (ABC) monitoring station and aerosol optical depth is identified as a key parameter to be observed. Linking with international research institutions to derive satellite-based information and building a regional knowledge hub will be important steps for understanding climate change. Such linking can be done by implementing the approach described here.

6. CONCLUSIONS AND FUTURE WORK

The SATOPI project demonstrates how Topic Maps technology can be used in order to federate Earth Observation data relevant to environment research with domain knowledge. An interesting characteristic of this standard approach is that each such application can be further integrated with similar applications thus providing interoperability among heterogeneous data sources. Especially in the environment field, this is a great advantage, as many phenomena are interlinked: earthquakes can cause floods or forest fires for example. The need of such integration is clearly endorsed by initiatives such as SEIS.

At the start of 2009, the work on an FP7 project – ULISSE – has begun. ULISSE which addressed the FP7 call “SPA.2007.2.1.01 Space Science”, puts as a target to pursue the exploitation and valorisation of scientific data from previous and future space science experiments on ISS as well as data from other space platforms. Although not related to environment data directly, ULISSE will benefit from the main ideas presented in this paper nut applied in a much larger scale: data provided from ten different scientific domains will be organized together with related domains knowledge using Topic Maps technology.

The authors of this paper believe that the lessons learned from these kinds of projects can be valuable to initiatives such as SISE and SEIS.

ACKNOWLEDGEMENTS

This paper describes the work done in the SATOPI project which is being developed by Space Applications Services as a co-funded activity with the European Space Agency (ESA contract number 21520/08/I/OL).

REFERENCES

Bajracharya S. R., Mool P. K., Shrestha B.R., *Impact of Climate Change on Himalayan Glaciers and Glacial Lakes. Case Studies on GLOF and Associated Hazards in Nepal and Bhutan*. ICIMOD/UNEP, June 2007, ISBN 978 92 9115 032 8.

Garshol L. M., Moore G., *JTC1 / SC34, ISO/IEC 13250-2: Topic Maps — Data Model*.

Mool P. K., Bajracharya S. R., Joshi S. P., *Inventory of Glaciers, Glacial Lakes and Glacial Lake Outburst Floods. Monitoring and Early Warning Systems in the Hindu Kush-Himalayan Region – Nepal*. ICIMOD/UNEP, August 2001, ISBN 92 9115 359 1, ISBN 92 9115 331 1 (CD).

Pinchuk R., Sharon Y.: *The Skin Design Pattern*. PLoP™2000, Allerton Park, Monticello, Illinois, USA, 2000.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

REAL - Remote sensing identification and monitoring of abandoned land

T. Soukup, L. Brodský, V.Vobora

GISAT s.r.o. - Charkovska 7, 101 00 Praha 10, Czech Republic (tomas.soukup@gisat.cz, lukas.brodsky@gisat.cz, vaclav.vobora@gisat.cz)

Keywords: Environmental monitoring, remote sensing, satellite Earth observation, land cover, abandoned land

1. INTRODUCTION

The process of land abandonment has become widespread in the last decades in all CEEC reflecting substantial political and economical changes in this period. Nowadays, it represent a serious issue and if not solved in time, the abandonment of productive agricultural lands would further grow. These lands would then gradually degrade together with their cultural landscape, soils will lose their productive potential, and jobs would be lost with consequent further migrations and further land degradation in a vicious circle that would deepen the social and economic problems of these regions. Besides the major social and economical impact, the process of land abandonment has also serious ecological consequences. It endangers valuable habitats to get degraded. Particularly biodiversity of semi-natural grasslands seems to be mostly affected by land abandonment [Baldock and Tar, 2002]. The biological consequences of land abandonment are discussed [Klimes, Jongepierova and Jongepier, 2000] analyzing long-term trials on rich species meadows in White Carpathians.

2. POLICY FRAMEWORK

Sustainable use and management of abandoned areas and prevention of abandonment land area enlargement is one of the policy targets both on European (e.g. Agro-Environmental Regulation, Council Regulation No 2078/92, Rural Development Regulation [EEC, 1992], Council Regulation No 1257/99 [EEC, 1999]) and national level (e.g. Biodiversity Strategy of Czech Republic (MŽP,2005), Post-accession Rural Policy 2004 – 2013 (MZE, 2005)).

Nevertheless, the assessment of the policy strategies requires clear measures based on the figures on evolution of actual abandoned land area. Unfortunately, there are currently no actual statistical data available on the extent of land abandonment so only rough estimates are used (e.g. total area of abandoned land about 300 000 ha, that is 7 percent of the total agricultural area (Ministry of Agriculture, 2001).

3. REMOTE SENSING POTENTIAL FOR ABANDONED LAND MONITORING

Remote Sensing is defined as the science and technology by which the characteristics of objects of interest can be identified, measured or analyzed the characteristics without direct contact. The characteristics of an object can be determined, using reflected or emitted electro-magnetic radiation, from the object. In general, each object has unique and different characteristics of reflection or emission based on the internal structure of object and/or its environmental condition. Remote sensing is a technology to identify and understand the object or the environmental condition through the uniqueness of the reflection or emission. Spectral reflectance is assumed to be different with respect to the type of land cover. This is the principle that allows the identification of land covers with remote sensing by observing the spectral reflectance or spectral radiance from a distance far removed from the surface.

Land cover mapping is one of the most important and typical applications of remote sensing data. Land cover corresponds to the physical condition of the ground surface, for example, forest, grassland, concrete pavement etc., while land use reflects human activities such as the use of the land, for example, industrial zones, residential zones, agricultural fields etc. Generally land cover does not coincide with land use and in general a land use class is composed of several land covers. Remote sensing data can provide land cover information rather than land use information, unless land use is clearly reflected by distinct land cover. From this point of view abandoned land represents more land use category with heterogeneous spectral reflectance characteristics as it is composed by mixture of grassland/pastures and shrubs colonized by herbaceous and bushy vegetation in various stages of succession (by various land cover). More, particular species composition also reflects specific topographic, climatic and pedology conditions of the locality.

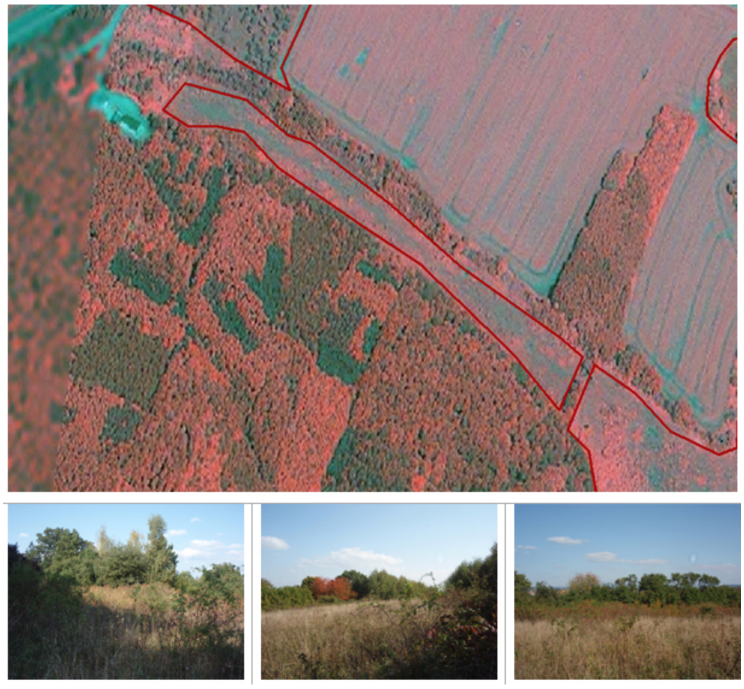


Figure 1. Abandoned land example as seen on VHR satellite imagery and on-site photos.

There has been number of land cover monitoring activities addressing the European level (e.g. LACOAST, MOLAND, CORINE LC, GSE Land, FTSPs), nowadays feeding the Land Monitoring Core Services (LMCS) concept development, and even more serving national and local users. Recent technological development together with growing demand for fast provision of land cover data updates leads to introduction of semi-automatic image processing methods and the integration of new technologies for land cover monitoring. Nevertheless, due to a complex nature of areas of abandoned land, this category is not present in recent nomenclatures (CORINE LC Methodology Guide, 2001, GSE Land Class Manual, 2006). This project aims to bridge this gap, explore the EO data potential and develop classification methodology for abandoned land identification, mapping and monitoring from the conceptual level down to the implementation.

Due to described high spectral variability in abandoned land areas, identification, mapping and monitoring by means of semi-automatic classification of remote sensing imagery is rather difficult task with number of fundamental research issues present both at the level of abandoned land definition specification, definition of class classification model as well as practical implementation of such model. Nevertheless, recent experiences with visual interpretation of abandoned land (DAIFOR, 2006)] shows, that use of recent advanced methodology could lead to feasible solution.

4. REAL PROJECT

The project aims to explore, develop and test new methodology for identification and monitoring of abandoned land via the state-of-the-art remote sensing techniques. Pilot areas have been used for demonstration of technically viable solution using object-based classification combining different strategies for abandoned land monitoring: spectral, textural, temporal and contextual ones. Methodology development was done in the general framework of Object Bases Image Analysis (OBIA) using the state-of-the-art DEFINIENS Enterprise software platform [Brodsky and Soukup, 2007a, 2007b] [Brodsky, Jupova and Kucera, 2007].

As already mentioned, abandoned land doesn't represent a single category of land cover and is more associated with changes in land use. These changes are then subsequently (and gradually) reflected in the landscape cover and its changes. From cultivated and cultural landscape over time it becomes a landscape of abandonment.

Similarly ambiguous is the definition of abandoned land itself. Most of the definitions which can be found in the literature are related to the legislative need for evidence of ownership, taxation purposes or the perspective of eligibility of subsidies. For example, FAO defines abandoned land "... as unused for agricultural production or other agricultural purposes for at least the next 5 years" [FAO's World Census of Agriculture (FAO, 1986, 1995, FAO / UNEP, 1999)). In contrast, the abandoned land as defined for the purposes of the common agricultural policy and agri-environmental measures support (Article 2.1 of Regulation (EEC) No 2078/92) refers to agricultural land, ".. that may be considered abandoned if it has not been the subject of an agricultural use or economic activity after a period of three years.

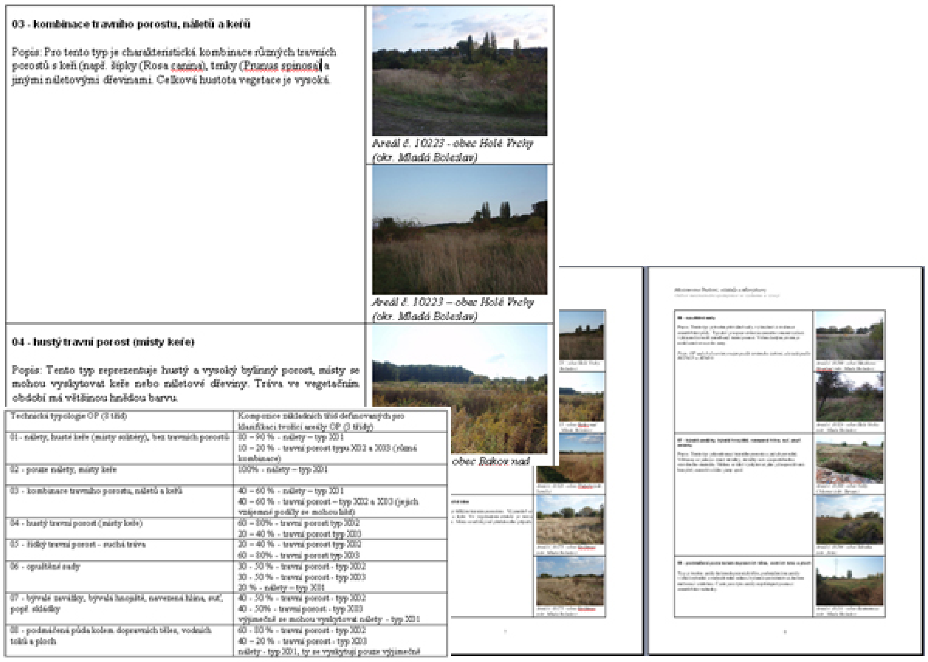


Figure 2. Abandoned land typology defined in the project includes 7 abandoned land sub-types based on various share and composition of 3 base "forming" elementary classes.

For our purposes, we therefore selected more technical definition of abandoned land from the perspective of the possible identification using remote sensing methods. Abandoned land, or more potentially abandoned land, is in our concept *a heterogeneous areas* *dominated by grasses, herbs and bushy communities in various stages of succession, with no evidence of agricultural use, or economic activity with a significantly different manifestations and evolution in the context of the surrounding cultural landscape.* Also a typology of abandoned land in the various sub-categories that define the type of structure and composition of abandoned sites has been proposed and used for abandoned land decomposition to types and base classes for classification purposes. 7 abandoned land sub-types has been defined in the REAL project based on various share and composition of 3 base classes (as seen on figure 2.).

Main activities in the REAL project can be summarizes as

- creation of knowledge base for abandoned land - the preparation, processing and implementation of the knowledge base in a form useful for flexible exploration and assessment of various characteristics of abandoned land on EO data, hypothesis evaluation etc.
- deep understanding of individual signatures, their quantitative description and selection of signature subsets to be used in classification model
- making the classification model for abandoned land classification using OBIA approach (implementation in DEFINIENS software)
- producing classification test on a test areas including validation

Abandoned land knowledge base

One of the major objectives of the project was creation of knowledge base focusing on abandoned land in terms of all defined strategies for identification of abandoned land on EO data - spectral, temporal, textural and contextual. Information obtained from the processing of remote sensing data and field measurement has been arranged in the form of a digital knowledge base - the library containing selected signature space for each abandoned land test site. The knowledge DB was implemented in the PostgreSQL relational database with the POSTGIS extension. The Knowledge DB can be accessed either directly through DB or GIS connectivity, but for a quick "exploring" the content of DB is also possible to access via web interface. This simple interface is created in PHP language, and enables fast and flexible browsing capabilities, according to the entered search criteria, hypothesis testing, etc. Open web client is also designed in a way that can be further modified or extended to new locations, for new signatures, but also to a new functionality in terms of output definitions and user functions.

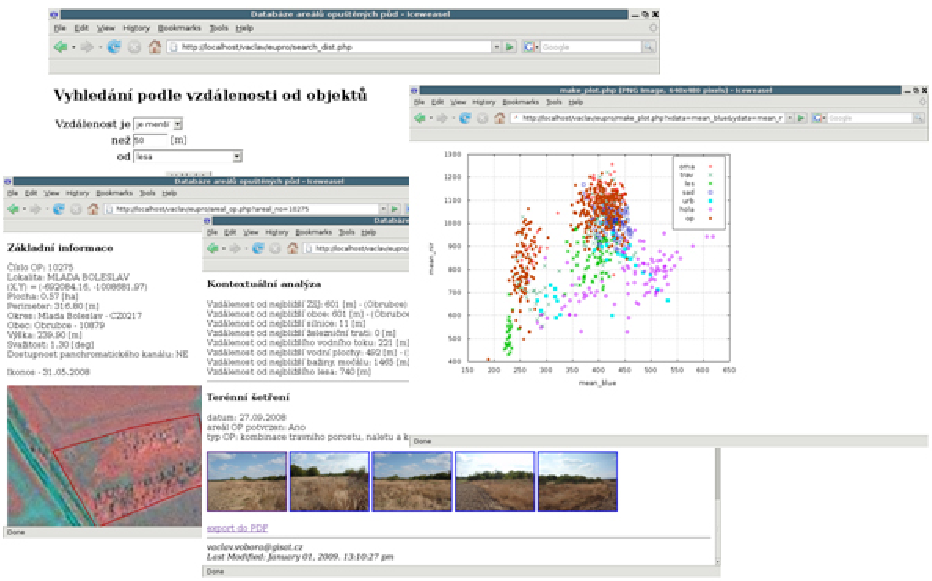


Figure 3. The knowledge DB for abandoned land implemented in the PostgreSQL relational database with the POSTGIS extension – examples of web interface

OBIA classification model

The resulting knowledge base was designed to support other main objective - making classification model and testing abandoned land classification. It allows quick and efficient verification of various hypotheses of classification rules for object oriented classification model in DEFINIENS. The creation of a knowledge base (Library) describing the symptoms of abandoned land on EO data represented a solid base for effective use and combination of automatic procedures and support the classification rules definition based on the use of fuzzy logic (membership functions and logical operators) and construction of a hierarchical structures (the rules if ... then). The aim was to build a flexible (transferable) model for classification of abandoned land objects based on the transparent

- descriptive knowledge (description of objects) based on designed and evaluated customized features of their transferability is assured by statistical analysis and
- procedural knowledge base featuring semantic grouping and contextual aggregation with defined parameters (aggregating of primary classes to heterogeneous super-classes based on class specifications) which is in particular applicable for abandoned land areas.

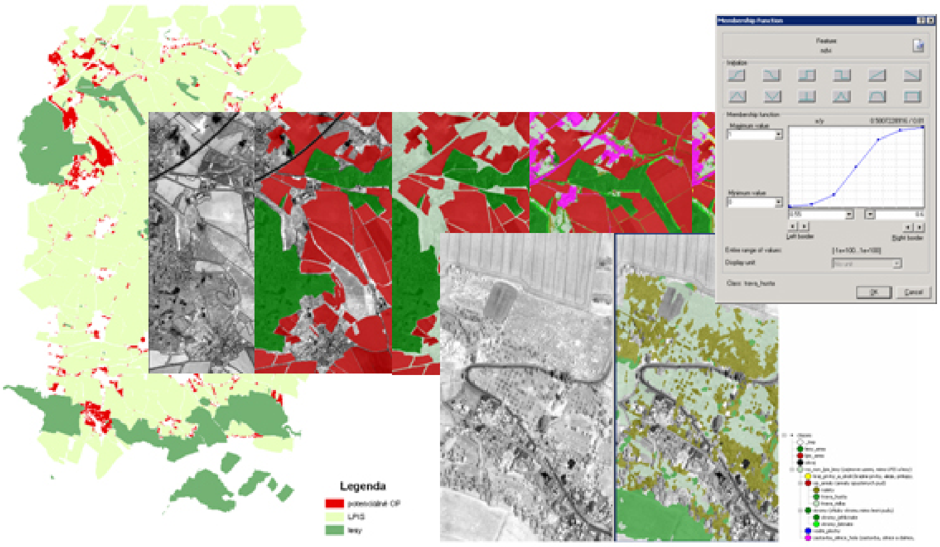


Figure 4. OBIA based classification model – classification hierarchy examples

Test classification and validation

Subsequently, the classification model was tested for real classification on test areas and verified using control areas of abandoned land within the test sites, including an evaluation of its accuracy. Three pilot areas were used for demonstration of technically viable solution using object-based classification. Testing of the classification was done on VHR and VHR/HR combination of images. HR data were tested to enrich spectral and temporal extent, however, HR data, with respect to abandoned land, are totally inadequate for use of textural rates. Overall accuracy moves relatively high (~80%, Kappa Index of Agreement *(*KIA*)* around 0.70), however, results clearly demonstrated also considerable variability for different types of abandoned land from defined typology. The result of our classification is twofold – identification of potential sites of abandoned land, defined as elementary base classes of abandoned land subtypes, but also the degree of association to the abandoned land class defined. Combining these results can be used to identify hot-spots abandoned land in the countryside.

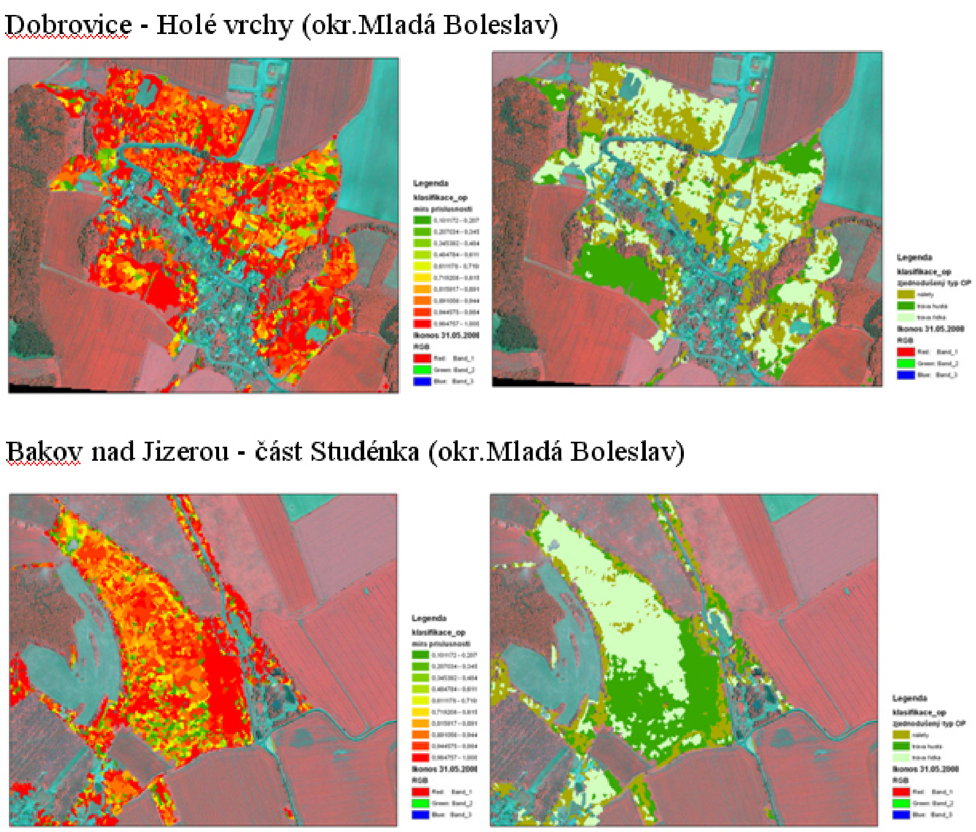


Figure 5. Example of abandoned land classification - identification of potential sites of abandoned land: degree of association to the abandoned land class defined (on left), elementary base classes of abandoned land subtypes (on right).

5. CONCLUSION & OUTLOOK

Land abandonment represents an important landscape change process with potentially substantial ecological, economical, social (and also political) consequences, while detailed up-to-date description and extent monitoring of this process is still missing. The REAL project - *Remote sensing identification and monitoring of abandoned land*, progressed to explore potential of the state-of-the-art remote sensing techniques for monitoring of land abandonment and developed and tested methodology based on OBIA approach for potential abandoned land identification. Three pilot areas were used for demonstration of technically viable solution using object-based classification combining three basic strategies for abandoned land monitoring: spectral, textural, temporal and contextual. Main outputs of the project include interactive web-based knowledge base for exploration of EO characteristic of abandoned land backed by PostgreSQL/PostGIS and classification rule base ready to be used in DEFINIENS software. Results, concepts and methodology developed, of the project are already further explored in similar activities run by GISAT at national and European level for identification of the vacant land in general.

ACKNOWLEDGEMENTS

The project was performed under the support of the EUPRO programme of Ministry of Education of the Czech Republic in the frame of the ERA-STAR GMES Pilot Project “Brownfields”.

REFERENCES

Baldock, D. and Tar, F. Agri-environment issues in Central and Eastern Europe – some implications of CAP enlargement, IEEP paper, London, 2002.

Brodský L., Soukup T. - Objektově orientovaná fuzzy klasifikace krajinného pokryvu (land cover) se zaměřením na zemědělskou krajinu, GIS Ostrava - DPZ a fotogrammetrie, Ostrava 29/1/2007 (a)

Brodsky L., Soukup T. - Object-Based Image Analysis For Land Cover / Use Classification, XII International Scientific-Technical Symposium - Geoinformation Monitoring Of Environment, Alushta/Crimea 10/9/2007 (b)

Brodsky L., Jupova K., Kucera L. - Semi-automated extraction of Landscape Features from VHR satellite data, Geomatics in support of the CAP, Palacio de Congresos, Madrid 12/11/2007

EEC - Council Regulation (EEC) - No 2078/92 of 30 June 1992 on agricultural production methods compatible with the requirements of the protection of the environment and maintenance of the countryside, OJ L 215/85, 30.6.1992, 1992

EEC - Council Regulation (EEC) - No 1257/99 of 17 May 1999 on support for rural development from the European Agricultural Guidance and guarantee fund (EAGGF), OJ L 160, 26.6.1999, 1999

Klimes, L.; Jongepierova, I. and Jongepier, J.W. - The effect of mowing on a previously abandoned meadow: a ten-year experiment. Praha: Priroda, 2000.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
Danielle Tacyniak, Agnès Marty
Thales Communications France, 2009

A Shared Environmental Information System for GMES: key topics

Danielle Tacyniak, Agnès Marty
Thales Communications, 1-5 Avenue Carnot, 91 883 Massy Cedex, France
(danielle.tacyniak@fr.thalesgroup.com, agnes.marty@fr.thalesgroup.com)

Abstract: The abstract describes some of the key topics which have to be solved mainly in the in-situ domain to target a sustainable and operational shared and distributed system enabling public services. It provides a description of the state of the art architecture OSIRIS and of the efforts to provide in order to have fully operational solutions.

Keywords*:* Environmental information, Sensor Web Architecture, sustainability, availability, responsibility.

1. INTRODUCTION

SEIS is a collaborative initiative of the European Commission (EC) and the European Environment Agency (EEA) to establish together with the Member States an integrated and shared EU-wide environmental information system. SEIS is mainly addressing in-situ information gathering from local / regional / national institutions, information processing, and analysis to be delivered to policy makers in many environmental domains.
GMES is a collaborative initiative of EC and the European Space Agency (ESA) dedicated to the surveillance of the earth and to the delivery of environmental and security-related information to users. GMES is focused on the delivery of core services principally based on space observing systems. Five large domains are identified today: Land, Atmosphere, Emergency, Marine, Security.
SEIS and GMES are different from many aspects: observation systems, data and information accessibility (voluntary service for SEIS, direct access for GMES), services.
However to better monitor the earth, manage situations at risk and help decision makers at all levels, these two initiatives must converge and solve jointly common issues.
A lot of technical, legal, financial, organisational and business model issues have already been addressed trough EU projects, by domain.
Nevertheless there are still a lot of remaining aspects to be covered to target a sustainable and operational shared and distributed system enabling public services, mainly in the in-situ domain which requires still more important efforts to be organised (compared to the space one), as mentioned many times during the Lille Forum 2008.

2. OSIRIS BEST PRACTISES FOR ENVIRONMENTAL SUSTAINABILITY

As an example, in the framework of the OSIRIS FP6/project, several user communities will soon benefit from internationally standardized pre-operational web-services.
The OSIRIS project targets the definition, development and testing of services dedicated to surveillance and crisis management, thus aiming at significantly enhancing the overall efficiency of the related operations. OSIRIS provides a Service Oriented Architecture based on OGC – Open Geospatial Consortium - standards and delivering functions ranging from in-situ earth observation to user services. It implements, and distributes Sensor Web Enablement services, as open source software. It also allows information management as

close as possible to its source and information availability to users in an open and transparent way thanks to web technologies.
OSIRIS concept and architecture are developed and validated in the framework of GMES and GEOSS. Pre-operational OSIRIS solution have been deployed and experimented in live demonstrations end 2008-beginning 2009 in different European areas in France, Italy, Germany and Spain.
Thus OSIRIS, called the “cornerstone of the key integrated European Infrastructure for effective environment management”, can be considered as an important input in SEIS and GMES.

2.1 The OSIRIS Services.

The OSIRIS services offer:

- The management of various types of sensor and data accessible via a proxy service,
- A dynamical sensor management (plug in of sensors and detection of failures),
- Logical connections between one or more applications, using predefined traffic patterns and an applicative Quality of Services (QoS),
- The access to data in different modes:
    - o Direct access to collected data (Pull access mode),
    - o Publish / Subscribe access mode to data,
- The integration of any web services compliant either the OGC Web Services Common Specification or the W3C Web Services specification,
- Mechanisms to develop adaptors to any services handling data.

The services are categorized into three groups, according to the problematic addressed, as shown on the figure below.

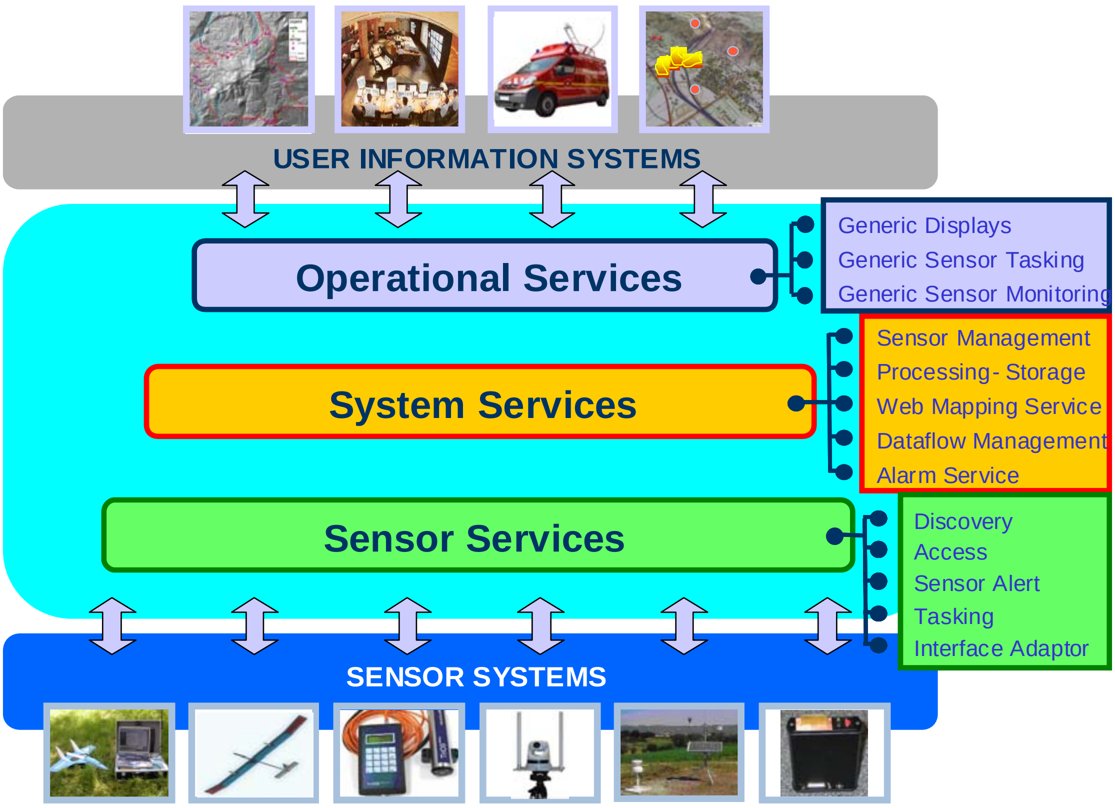


Figure 1. The OSIRIS architecture

From the bottom to the top, this corresponds to:

- **Sensor Services:** OSIRIS provides a set of services for accessing sensor data and controlling sensors. Standardization is a very important step in order to reach interoperability, thus those services are based on the SWE, adjusted for OSIRIS. The main improvements are done in conjunction with OGC, and are participating to the new revisions of the standards,
- **System services** providing additional capacities upon sensor services, better addressing a system level:
    - Data processing infrastructure, managing and orchestrating processing means. An easy and documented way to integrate processing means is provided,
    - Data Storage offering storage capabilities and access to the stored data, hiding the complexity of the underlying data sources involved.
    - Alerting means, allowing users to register then be notified according to their wishes and the events occurring. This service supports major alerting protocols, like protocols coming from OASIS, and is able to process several event sources, like SAS (Sensor Alerting Service) from OGC. Alerts can be sent through SMS or email, or through RSS that allows easy integration in a Web portal.
    - One dedicated set of services is implementing the necessary glue between the other services; it is called the System of System Service Layer composed of:
        - Registration of services participating to the system, and supervision of the services. The Sensor Instance Registry service is there coupled for providing the information on the services deployed at the sensor level.
        - Data Flow Organisation: using any data provider service (SWE or Web Service or other) as data source to feed the Data Storage for example.
        - Integration means for any service thanks to customisable and documented connectors, facilitating the integration of existing services.
- **Operational services**: addressing in particular the display of the sensors data and System Services information. Applications are delivered for supervision purpose, and exploitation of the services (data display, alert display, tasking of sensors for example). Those capacities are addressed in OSIRIS with a Generic Display and Web Display, and benefit from the choices made in the architecture, like the use of Web Services and the OGC standards. Some operational services are also dedicated to one specific domain, or to specific sensors.

The services are constituted in several layers, the sensor services (near the sensors) being exploited by added value services (system services). Operational Services can access the System Services as well as the Sensor Services.

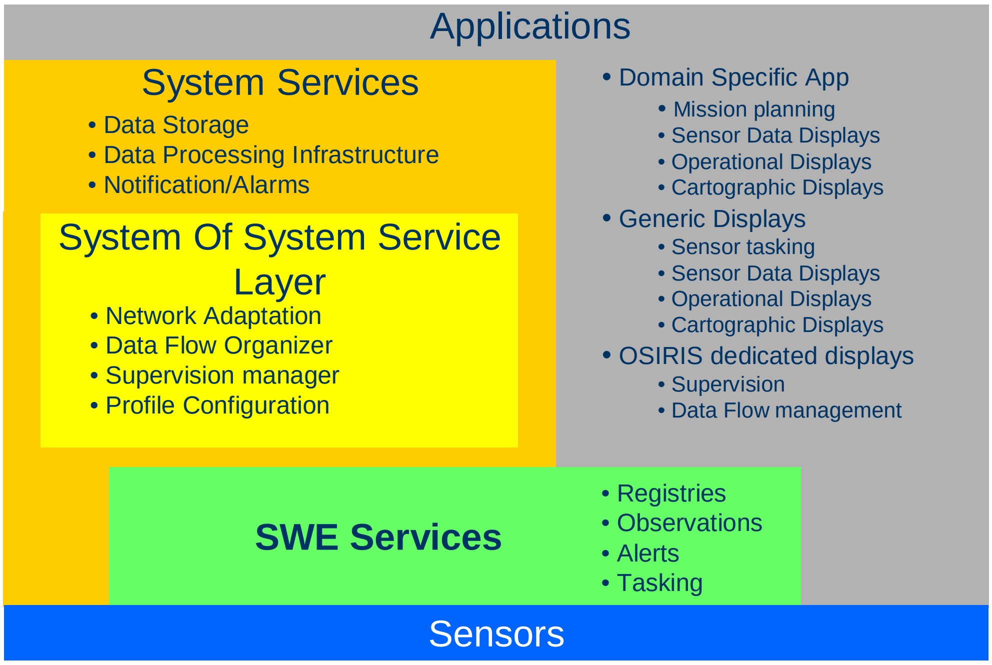


Figure 2. The services offered

All the services at sensor level or system level can be directly accessed either by internal or external applications. Security policies may restrict access to some services, according to the system deployed.

2.2 Deployments of the Services in a System

The OSIRIS architecture is made to be flexible. In order to build an operational system relying on the OSIRIS architecture, some services are mandatory whereas some other optional.

The OSIRIS services have been successfully demonstrated in four live experiments: Forest fires, fire in industrial buildings, fresh water management, air pollution, each covering the monitoring and crisis aspects, and complementary in terms of:

- Environmental constraints,
- Time constraints,
- Smart sensors based on fixed or mobile platforms, including airborne platforms especially a UAV,
- Information delivered by the sensors.

The demonstrations allowed to show in real conditions, the whole processing chain, from information capture to exploitation by the end-users.

Figure 3. The deployments

3. REMAINING ISSUES FOR OPERATIONAL DEVELOPMENT AND IMPLEMENTATION

3.1 From SEIS concept to reality: operational architecture and implementation road map

Starting from existing e-infrastructure, systems and services in Member States and EU institutions, from R&D projects such as OSIRIS, from SEIS and GMES concept and from European directives (INSPIRE, Water, …), objectives are to establish a target solution and a step by step implementation road map for the delivery of a decentralised but integrated web-enabled information system based on a network of public information providers sharing environmental data and information.

First step shall be the implementation of OSIRIS concept and solution to move away from a today "paper-based reporting" to a system where information is managed as close as possible to its source and made available to users in an open and transparent way. This operational implementation will enable gathering of in-situ data and management of observing sensors through a sensor web-enablement, electronic storage of environmentally-related data, and accessibility of information to different user communities through web services.

3.2 Business models

In the framework of GMES, inventory of EO existing systems as well as targeted operational and business model solutions are or will be analysed and defined for each of the 5 large domains, independently.

Business models for the in-situ market have not yet been established.

One solution which could be studied is based on the concept of "Data Public Stock Exchange". Objective of this service is to encourage data/information sharing in different public entities at regional level and to raise European capitalisation for the benefit of public sectors as well as the private ones.

3.3 Legal and security aspects – Data policy

Information Security is of most importance in the EO community:

- INSPIRE directive and willingness of European countries to implement it is essential but not enough
- Co-operations at local, regional, European and international levels have to be clearly defined in terms of data/info/services to be exchanged
- Information Security covers different aspects such as reliability, confidentiality, which are common to all GMES domains (Land, Atmosphere, Emergency, Marine, Security), however specific aspects have to be analysed for the Security domain
- Individual vs aggregated information confidentiality has to been tackled, mainly in case of free access to citizens
- Undermining civil liberties has to be taken in consideration (public access to private EO info)

Until now information security actions have been focused on:

- End to End data security concept study funded by ESA (space component)
- Information security requirements (ISO norm) for core services (limited to willingness to implement information security processes)

It’s why we recommend to launch a study on data policy in GMES:

- to analyse political, juridical, technical aspects regarding Information Security in GMES
- to establish global and specific recommendations for data policy in GMES, in Europe and outside Europe in the framework of international co-operations.

ACKNOWLEDGMENTS

The authors wish to thank the European Commission for the financial support of the FP6-OSIRIS project.
www.osiris-fp6.eu

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I. Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Czech approach towards climate change

Jan Pretel, Radim Tolasz
Czech Hydrometeorological Institute, Na Šabatce 17, Praha 4, Czech Republic
tolasz@chmi.cz
pretel@chmi.cz

Actual climate trends in Czechia are posted to world and Europe climate context. The development of past climate in Czechia has a good conformity to European climate change as presented in AR4 IPCC report. The same is valid for projections of future climate trends for our region in comparison with European continent.

The Czech climatologists are running the ALADIN-CLIMATE/CZ model. Result of projections for SRES A1B seems to be very useful and applicable. Preliminary simulations are made for temperature and precipitation changes to the year 2030.

Simulated changes of temperature (°C) for year 2030 (reference period 1961-1990)

percentile / season	**MAM**	**JJA**	**SON**	**DJF**	**year**
Q50	1,1	1,0	1,3	1,4	1,2
Q25	*0,7*	*0,8*	*1,1*	*0,9*	*0,9*
Q75	*1,5*	*1,5*	*1,6*	*1,8*	*1,6*

Simulated changes of precipitation for year 2030 (% of reference period 1961-1990)

percentile / season	**MAM**	**JJA**	**SON**	**DJF**	**year**
Q50	7	3	1	4	4
Q25	*-1*	*-4*	*-5*	*-1*	*-3*
Q75	*12*	*8*	*8*	*7*	*9*

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Better geoinformation for spatial planning: Implementation of GMES services in the Czech Republic

J. Weichselbaum[a], **T. Soukup**[b]

[a] *Geoville Information Systems Gmbh, Museumstrasse 9-11, 6020 Innsbruck, Austria (info@geoville.com)*

[b] *GISAT s.r.o.,Charkovská 7,101 00 Praha 10, Czech Republic (gisat@gisat.cz)*

Keywords: GMES, Spatial Planning, Climate change, Urban growth, Earth Observation, Soil sealing.

1. INTRODUCTION

Integrated spatial planning is a key element to limit human and socio-economic loss, where, for example, land take by urban growth may lead to increased flood impact through faster run-off and higher exposure of population. In response to reporting obligations arising from a series of environmental directives and policies and for mitigation of climate change impacts, the Spatial Planning Service acts as one of the core information services in the land domain of GMES (Global Monitoring for Environment and Security).

2. GMES SPATIAL PLANNING SERVICES FOR PUBLIC ADMINISTRATIONS

The Spatial Planning Information Services are based on the combination, analysis and modelling of data received from Earth Observation satellites as well as in-situ measuring networks and statistical data. Thereby this service provides cost-effective, wide-area and cross-boarder harmonized geo-information products for a multitude of territorial applications, helping to improve the quality of life of the European citizens through a better protection of their environment, the enhancement of their living conditions and the reduction of risk potentials and threats.

The spectrum of users of the service comprises spatial planning departments within local, regional governments as well as Federal Environment Agencies and national ministerial bodies. In addition, the service caters to the specific information needs of European Commission Services.

The *Spatial Planning service* offers a solution to monitor urban growth and soil sealing from European down to regional or local administrative levels and provides insight into recent and future transboundary land take trends and their impact on the environment. The service offers a portfolio of spatially referenced and consistent geo-information products (maps, statistics, indicators and scenarios), addressing specific information needs arising from a number of regional, national and European directives and policies, thereby supporting spatial planning authorities across Europe in their efforts to fulfil a broad range of monitoring and reporting obligations.

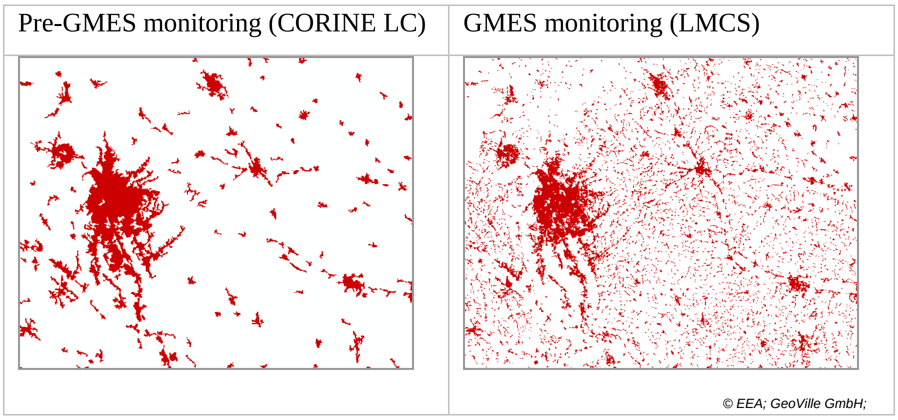


Figure 1. GMES reveals a ‘real-world’ picture of built-up areas, enabling regional governments and European authorities to engage in effective spatial planning and territorial monitoring based on more detailed information.

3. IMPLEMENTATION OF THE GMES SPATIAL PLANNING SERVICES IN THE CZECH REPUBLIC

In the frame of a series of GMES projects funded by the European Commission[1] and the European Space Agency[2], several bodies of public administration have been already served in the Czech Republic. Within the ESA funded GSE SAGE, the Spatial Planning service has been successfully implemented and validated by the Pilsen Regional Authorities and the Ministry of Environment for an area of 35.000 km² in the western part of the Czech Republic. Within the ongoing ESA funded GSELAND project (follow-up of the GSE SAGE) further spatial planning support services are served to the Czech users - Pilsen Regional Authorities, City Development Authority Prague, Moravia-Silesia Regional Authorities, Czech Ministry of Environment [Nalevka and Soukup, 2007].

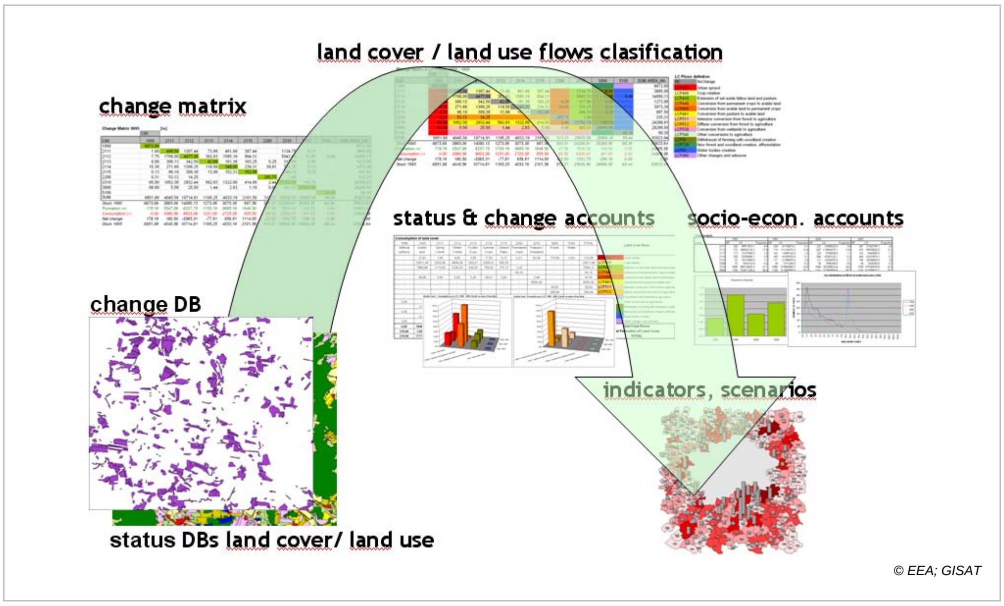


Figure 2. Spatial planning services - general concept: from data to spatial insight, assessments and scenarios modelling

[1] FP6 Integrated project geoland and the recently started successive FP7 Collaborative project geoland2 (http://www.gmes-geoland.info)

[2] ESA GMES Service Elements (GSE) SAGE (Service for the Provision of Advanced Geoinformation on Environmental Pressure and State, http://www.gmes-sage.info) and GSE Land (http://www.gmes-gseland.info)

The Spatial Planning service provides spatially referenced and consistent geo-information to describe, explain and forecast urban land take and the impact of land consumption on the population and the environment. The service allows for monitoring of urban structure, urban growth and soil sealing, and describes the pressure, state and impact of urban land take on regional and European scale – see example of concept figure 2.

The *Spatial Planning geo-information* services are based on general geo-information on land cover and land use derived from high-resolution Earth Observation (EO) data. Although, there is certainly a heritage of European-wide land monitoring projects (e.g. CORINE Land Cover), within GMES, European service providers have developed highly-automated land use mapping applications and joint processing chains that further increase the reliability and data throughput of traditional processing

A good example of strength and effectivity of such joint service production approach was the cooperation of GISAT and GeoVille companies for the Czech Republic users. The processing chain for automated land use mapping was based on Definiens Cognition Network Technology® - robust context-based technology designed to model the human cognitive perception processes to extract information from images. Utilization of the Definiens software was essential to the classification as the enterprise architecture allowed to develop solution that can be easily adjusted in the production process providing faster image analysis and deeper insights into the problem. This solution was also awarded by Definiens (Definiens European GMES Innovation Award for GeoVille). The result provides user with harmonised, cross-border land use data that enables users to make better decisions [Brodsky and Soukup, 2006, 2007], [Weichselbaum, 2007], [Weichselbaum, Hoffmann, Kleeschulte, 2007].

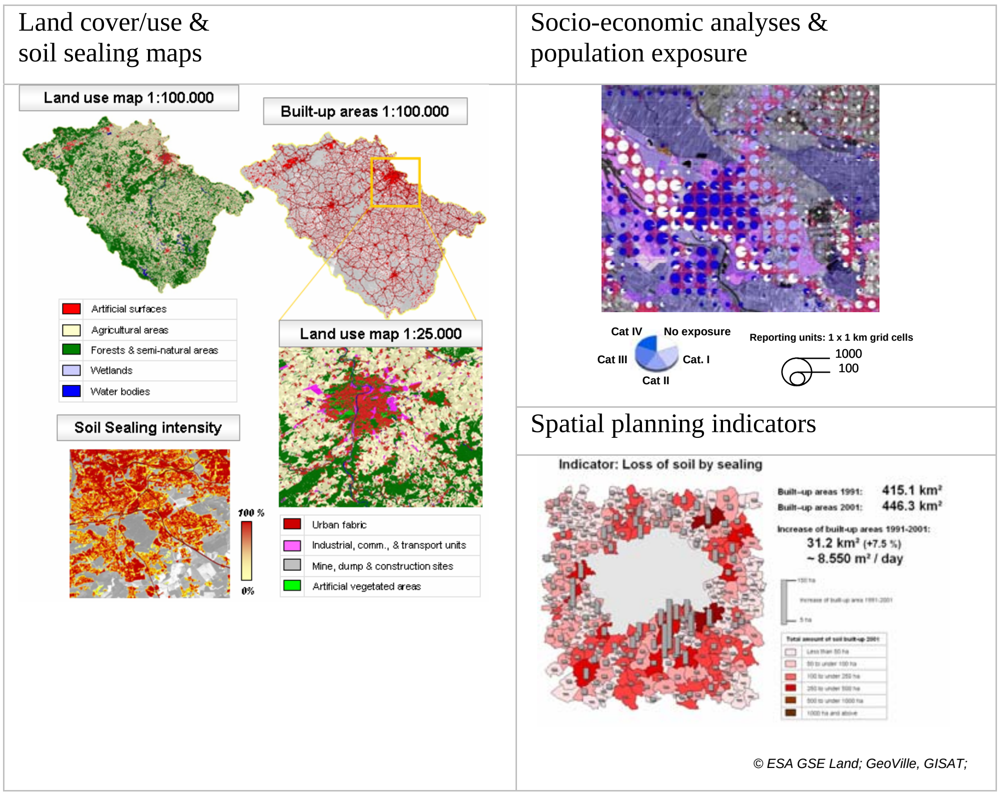


Figure 3. Integrated product portfolio for spatial planning, bringing together Earth Observation and in-situ data – example Pilsen region and Prague.

The land cover/land use maps depict with great accuracy the extent, development and density of built-up areas (Figure 3, left image, Figure 4, left). They are subsequently integrated with ancillary geospatial and statistical data into Geographical Information procedures and models. These derived information products open the way for analysing demographic developments and urban land take trends and for describing the state of land consumption

Land cover/use & soil sealing maps	Socio-economic analyses & indicators Spatial insight into the area development

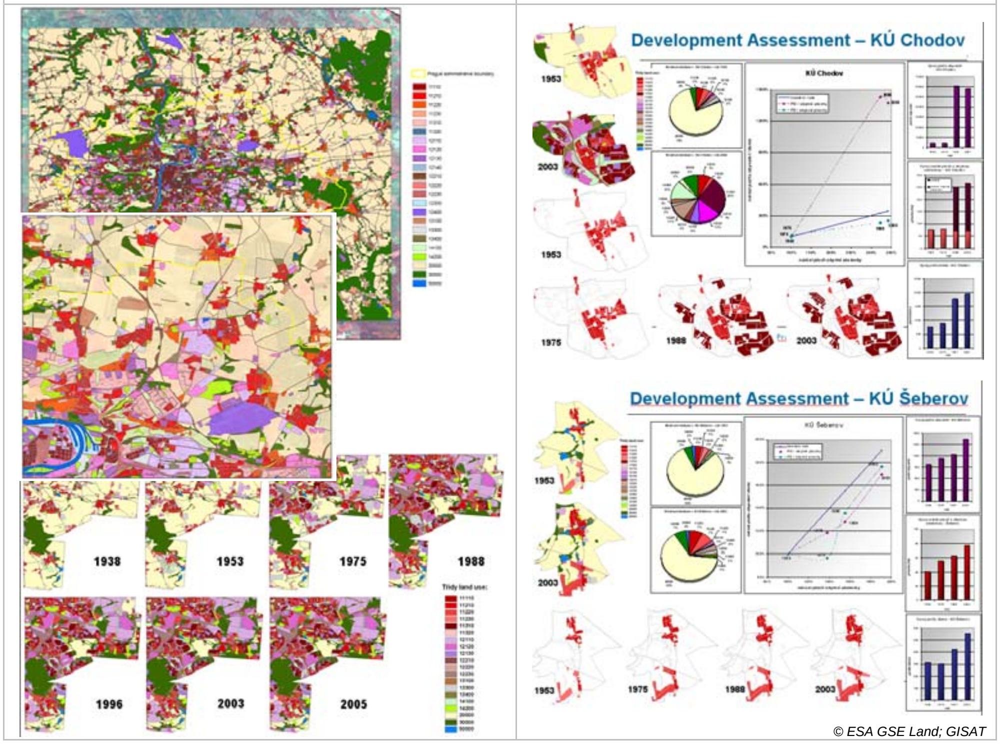


Figure 4. Integrated product portfolio for spatial planning, bringing together Earth Observation and in-situ data – example for Prague area (Urban Atlas).

and its impact on the environment. An example is the utilisation of the spatially explicit built-up areas to refine coarse demographic statistics by spatial disaggregation modelling [Steinnocher, Weichselbaum, Köstl, 2007]. The resulting localised statistics are a key input for population exposure analyses and risk management applications (Figure 2, upper right image) and spatial planning indicators on local, regional or European level (Figure 3, lower right image, Figure 4, right). As a further element of the product portfolio, urban growth models can be employed to simulate and forecast potential scenarios of future urban land use changes.

Tab. 1: Key policy questions addressed by GMES Spatial Planning services and related service benefits

Selected key policy questions answered	Service benefits
▪ How much and in what proportions is land being taken for urban and other development? ▪ Where does the most significant land-take occur? ▪ What are the drivers of uptake for urban and other artificial land development? ▪ How many people are affected?	▪ Enriching lump statistics with geospatially explicit information ▪ Guaranteeing European consistency and comparability ▪ Facilitating the evaluation of policy options ▪ Improving décision-making through better planning information ▪ Moving from observing and monitoring to policy evaluation

4. CONCLUSION AND OUTLOOK

SAGE and GSE Land projects have implemented Spatial Planning services in regions across 9 European countries. It has also served as an important reference for the first implementation step of the European GMES Land Monitoring Core Service via the "GMES fast track service on land monitoring - High resolution core land cover data for built-up areas, including degree of soil sealing 2006" [Weichselbaum, Gangkofner, Hoffmann, 2007], implemented by the European Environment Agency together with the European Commission and the European Space Agency. This first roll out of the service covers 38 countries spanning from Portugal to Turkey and from Sicily to the North Cape. Also recent Urban Atlas activities represent another achievement of the GSELAND project, where the Urban Atlas service has been designed and tested. The Urban Atlas is currently financed by EC with the support from the European Regional Development Fund (ERDF) and full operability is foreseen for 2011. The first edition of the Urban Atlas will be delivered in 2009 to 185 cities

The benefits of the provided products for public entities on local, regional and European level have been successfully evaluated and are manifold. By using advanced image processing algorithms and GIS modelling techniques, the service provides detailed, homogeneous and up-to-date land use maps and derived information products. The service contributes to improved decision making by enriching and sharpening lump statistics with spatially explicit information. Furthermore, the products are more detailed with regard to thematic class definition, scale and geometric precision than the currently applied data sets (e.g. CLC, cadastre, land use zoning plans and statistical land use information). The service delivers GIS-ready geo-information products, which can be directly ingested in user-side expert information and decision-making systems. Also, this service framework set-up perfectly fits to the overall European-wide initiatives related to standardization and data provision framework, represented by INSPIRE and SEIS, as the detailed, high-quality data guarantees consistency and data comparability from regional to a European level and allows the development from observation and monitoring of environmental trends towards the evaluation of policy options.

Besides the LMCS components which has already reached certain level of operational status (e.g. Urban-Soil Sealing or Urban Atlas, as mentioned above), the spatial planning services provision to the Czech users will continue under the recently started EC-funded FP7 project geoland2. The Czech Users, under umbrella of Environmental Information Agency CENIA, are involved within further the Spatial Planning Core Information Service developments. The full range of harmonised EO based products and tools to describe, explain and forecast urban land use changes in the Trans-European Network (TEN) corridor Dresden-Prague-Linz will be operationally demonstrated, implemented and validated until 2012. The aim is to further operationalize the Spatial Planning Information services towards one of the elementary information services inside of the operational GMES Land Monitoring Core Service (LMCS).

REFERENCES

Brodsky, L. Soukup, T. (2006): Development of agro-envi indicators based on land use/land cover changes for assessment of rural landscape changes. Geographical information in support of the CAP, Toulouse 27 – 29 Nov. 2006

Brodsky, L., Soukup, T. (2007): Object-Image Analysis For Land Cover/Use Classification, Proceedings of XII. International Scientific-Technical Symposium Geoinformation Monitoring Of Environment: GPS And GIS Technologies September 10-15, 2007, Alushta, Ukraine

Nalevka, O. – Soukup, T.: GMES – Remote sensing for public administration and self government in the Czech republic. 1st National Congress GIS - Geoinformatics for Everyone, Mikulov, 2007. ISBN 978-80-86633-79-4

Steinnocher K., Weichselbaum J., Köstl M. (2006): Linking remote sensing and demographic analysis in urbanised areas. In: Hostert, Damm & Schiefer (Hrsg.): First Workshop of the EARSeL SIG on Urban Remote Sensing "Challenges and Solutions", Berlin, 2.-3. März 2006, CD-Rom.

Weichselbaum, J. (2007): Intelligent Geoinformation to Go! ESRI ArcNews, Summer 2007. (Web: *http://www.esri.com/news/arcnews/summer07articles/intelligent-geoinformation.html*)

Weichselbaum, J., Hoffmann, C., Kleeschulte, S. (2007): Transforming Images Into Intelligent Geo-Information, 79 – 87 Proceedings LCL 2007 International CODATA Symposium on LandCover Logic Bonn, Germany, November 28/29, 2007.

Weichselbaum, J., Gangkofner, U., Hoffmann, C. (2007): GMES Fast Track Precursor delivers European high resolution information on built-up areas. Erdas User Group Meeting 2007.

Workshop 2

ICT-ENSURE: European ICT Environmental Sustainability Research

Organized by **Werner Pillmann**, Jiří Hřebíček and Antonios Barbas

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

WISE: Water Information System for Europe, issues and challenges for Member states

D. Biliouris[1], J. Van Orshoven[1]

[1] K.U.Leuven, Department of Earth and Environmental Sciences, Spatial Application Division (Dimitrios.Biliouris@sadl.kuleuven.be)

Abstract: In the framework of the WISE-initiative 'Water Information System for Europe', EU-Member States have to report a number of water related data. For the harmonized incorporation of this data into the WISE system, EU and EEA contracted several consultancies to perform diagnostic analysis and to provide a number of tools and services. The Spatial Applications Division of the University of Leuven (SADL) together with the Belgian company Geosolutions, has been acting as such a consultancy. An overview of the findings in terms of the issues and challenges the Member States are facing in order to comply with the WISE needs is highlighted. Two separate issues are described, namely the findings regarding the diagnostic analysis on ART.5 data and the procedure development for a Quality Analysis/ Quality Control (QA/QC) tool.

Keywords: WISE; Water framework directive, data reporting, EC-DG ENV; EEA; Member States

1. INTRODUCTION: WISE AND ITS CONTEXT

WISE ‘Water Information System for Europe’ is a shared information system providing water related information available on European level. It stands for modernising and streamlining the collection and dissemination of information related to European water policy. The formal reporting framework for WISE is the European Water Framework Directive (WFD). WISE is a joint activity of the European Environmental Agency (EEA), its 32 Member States organized in the European Environment Information and Observation Network (EIONET), the Joint Research Centre (JRC), Eurostat, and the European Commission (DG-Environment).
WISE provides background information on EU legislation, thematic knowledge on the European waters,and on the status and evolution of the water resources. Furthermore, it provides access to data centre services, reports, indicators as well as Europe covering maps and data access for diverse aspects of water. WISE is intended to be fully operational by 2010 (DG Env, 2008).

WISE requires water related information for a number of reasons, namely:

- To check compliance with the requirements of specific articles of the WFD;
- To carry out preliminary assessment of the situation in the Member States;
- To carry out further detailed analysis (where additional data may be required);
- To compile statistics for its own needs and to inform the European Parliament and
- To create a European-wide picture to inform the public.

Furthermore, the Water Framework Directive has adopted a number of important key elements and requirements to which the WISE system is to contribute, such as:

- Protection of all waters and achieving “good status” for all waters by 2015;
- Covering all impacts on waters (e.g. human activities);
- Development of Registers (e.g. protected areas, significant pressures, …);
- Water quality defined in terms of biology, chemistry and morphology - Monitoring of data and Data management;
- Co-ordination in International River Basins;
- Active Public Participation (Public information and consultation);
- River Basin Management Plans;
- Economical Aspects;
- Reporting Obligations.

WISE is an international community network that deals with an interactive data flow-in and flow-out among data providers and users with the fundamental input providers being the individual Member states that report all the necessary data (figure 1).

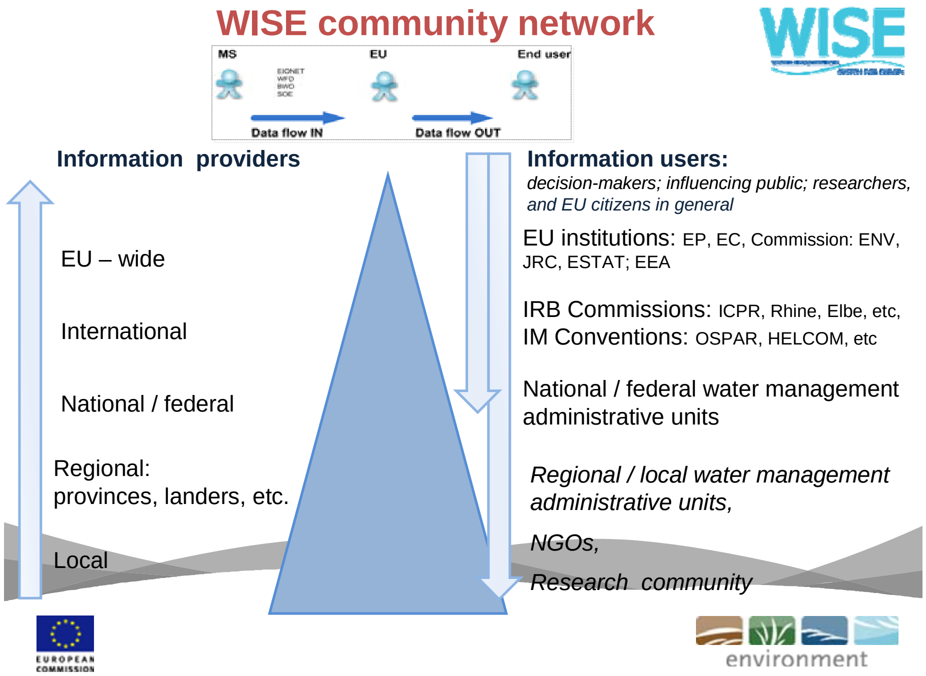


Figure1. WISE community network. Information providers and users. (taken from dr. Violeta Vinceviciene presentation in WISE workshop, PARIS 12/11/08)

Thus national reporting is the most important tool of the implementation of the EC Directives. On a national basis, reporting involves the following:

- The fulfilment of obligations of the EC Directives;
- Showing the success of taken measures;
- The basis for the Compliance Check for the European Commission;
- The basis for Synthesis reports of the European Commission;
- The basis for the evaluation of policy effectiveness of directives;
- The basis for the State of Environment analysis;
- The starting point of infringement procedures.

The short and long term objectives for the spatial data components of the WISE are outlined in the WFD guidance document on GIS and in the WISE implementation plan (Hannerz and Langaas, 2007; Cima et al., 2005; Common Implementation Strategy for the water framework directice (2000/60/EC).

2. WISE AND CONSULTANCIES

The overall objective of consultancies is to assist the ongoing development of strategies for the reporting on the European water related directives with concrete tools, services as well as guidance of the GIS-components. This encompasses the provision of services and development of tools in implementing WISE for reporting and data management on both existing and new or upcoming directives. SADL/Geosolutions is such a consultancy that provides a number of services under the WISE framework, mainly on diagnostic analysis of submitted data, development of QA/QC tools, implementation plan for a WISE distributed system and so on.

2.1 Diagnostic analysis on ART.5 data

SADL/Geosolutions performed a diagnostic analysis on the reported data, under the reporting obligations of Article 5 of the WFD, of each Member State on both geographic and alphanumeric inputs of the characteristics of their River Basin Districts (RBD). This diagnostic analysis focused on the way the data were reported and the compliance of the reported data with Art.5 requirements. As result per each MS a detailed report was provided containing:

- Missing Data;
- Obvious errors;
- Problems in data report that hinder further analysis;
- General remarks and
- Recommendations towards a harmonised data submission at EU level.

Since WISE is addressing Member States independently there is a difference on the submission per each Member state, while a number of Member States miss the proposed deadlines for a number of reasons.
On the submitted data there were obvious wrong locations in the geodatasets due to data input mistakes (e.g. Latitude decimal should have been shifted one number, neglectance of inputting the coordinates or just inputting the wrong positions).
Furthermore, regarding centroids of surface water bodies there was variability among the Member States in the way they were reported. Some countries delivered the geodatasets in geographic coordinates, others in projected coordinates and more than once the coordinate system was unknown.
In general within Member States data quality differs enormously and therefore a full inventory is a prerequisite to successful QA/QC procedure implementation and subsequent compliance checking.
Per each water theme and protected areas a number of quality assurance issues were stated concerning centroids, polygons, as wells as adjacency effects and location of water bodies within the boundaries of member states and River Basin Districts.
On the submitted data a number of checks were performed for all Member States, regarding:

- Centroids
- Coordinates
- River Network
- Connectivity
- River stretches
- Slivers
- Variability inside MS

The comparative analysis of the reported data followed several hierarchical "checking" steps such as:

- File naming;
- File/Folder organisation;
- Mandatory Fields (filled or not);
- Topological errors;
- Layer statistics;
- Referential Integrity;
- Positional accuracy;
- Fields reported in geometry files by the Member States;

On the basis of the diagnostic analysis of the files reported by the Member States, a number of recommendations were made concerning the organisational structure of files and file names, the overall quality improvement of data, and the quality of geographic data. The areas where significant improvements can be made for each Member state are listed below.

❖ ***File organisation***

Concerning file organisation, the following recommendations pertain:

1. Versioning of files should be avoided. Different versions should be incremental improvements that take into account all corrections made in previous files;
2. Cascading directories are to be avoided. This makes the structure of deliverables per Member States more transparent;
3. Standardised metadata are needed that describe *inter alia* the origin and corrections made;
4. File naming should be consistent so that automatic uploads are facilitated.

❖ ***Geometry files (shapefile format)***

The geometry files that were reported contained a huge variability in quality between the Member States. Specific recommendations were given, related to referential integrity, projection systems, accuracy and topology.

Referential integrity:

- Codes should match and also field names. This enables cross-checks between layers.

Projection system:

- The projection system should be defined. The recommendations of the GIS guidance are GCS_ETRS_89 as geographic coordinate system (GIS guidance, 2000/60/EC)

Accuracy:

- The recommendations in the GIS guidance are 125 meters with a minimum of 1000 meters.

Topology:

- Overlaps, overshoots, undershoots, slivers to be avoided;
- Connectivity of river network is desirable, certainly within the country.

❖ ***Data files (xml-format)***

Regarding xml files recommendations were provided in a similar fashion as for the shapefiles.
Referential integrity:

- Codes should match between the different files and should be according to the guidelines. This enables automatic cross-checks between layers. At present many differences exist in coding..

Accuracy of reporting:

- All field names in the predefined schemas should be present and used correctly;
- The records should be filled, including the Member State coding.

Considering the amount of input data on each reported schema by the MS, we concluded that¸ a thorough and rigid QA/QC tool should be proposed. This QA/QC tool should provide a constant validation on the submitted data. This validation check should focus on:

- Integrity and completeness of the data;
- File format harmonization;
- Data harmonization and
- Geometry harmonization.

2.2 Procedure development for QA/QC tool

The Water Information System for Europe provides a centrally-based repository for the submission, validation and dissemination of data reported under the Water Framework Directive. As well as enabling Member States to provide their submissions to meet the reporting requirements of the WFD in standard electronic formats, WISE also provides automatic and manual validation on the submitted data, before uploading into a centrally managed spatial database that enables the web-based query, view and analysis and mapping of the datasets.

WISE allows Member State users to submit data in several different formats, based on XML, GML and ESRI Shape files as appropriate. The formats are tightly defined to match the requirements as outlined in the reporting guidance documents. For each phase of reporting, the data requirements are defined as an XML schema for attribute/text, and GML/ESRI shapefiles for spatial/geographic data. Each individual data requirement is termed a ‘component’. For Article 3 data each component was provided at a Member State level. Article 5 data is provided at the individual River Basin District level.

The principal aim of this task was to provide the tools so as completeness and referential integrity between the different components of the ART5 reported data is secured. The different water bodies and protected areas are taken under consideration and assessed for accuracy reporting.

At the data input stage, QA/QC tools comprise a set of validation rules that enable a first assessment in terms of integrity and completeness of the data. At the data control stage, QC tools will automate a procedure or set of procedures intended to ensure that the information delivered adheres to a defined set of accuracy and comparability criteria.

The procedure development of the QA/QC is in accordance with EEA, the GIS guidance and reporting obligations of the Member States.

QA/QC tools should check and conclude on harmonization of the different types of reported data, such as XML schema, HTML templates, and GML or ESRI Shapefiles, at both national and international levels.

❖ ***Harmonisation***

Harmonisation has two aspects – File Harmonisation and Data harmonization. File harmonization considers the naming of the submitting files/data/schemas, while data harmonization deals with the harmonisation of the content (attributes, metadata) and harmonisation of geometry.

- File Harmonisation

Since the MS report their files independently, a QA/QC tool should be created that would allow WISE to check and harmonise all the different submitted files.

In order to avoid confusion with resubmission and different versions the two mandatory attributes required for each XML file such as the Creation Date and the Creator should be checked so as always the latest version should be considered. Specifically the versioning of files should be checked to account for added data. However, as proposed previously all the different versions should be incremental improvements taking into account all corrections made in previous files. Ideally transmission of multiple versions should be avoided. Furthermore the amount of existing directories should be checked to ensure that all the required data are included. Meanwhile, the naming of these directories should be checked for consistency so as an automatic upload procedure to be facilitated.

- Data Harmonisation

The QA/QC tools should follow the following validation on each of the submitted schemas::

1) Check Element Naming;
2) Check Mandatory fields;
3) Value domain restrictions. According to a list of acceptable values for each element a QA/QC tool should be present to ensure that only valid values are passed;
4) The File name and identification should be checked per schema and water body;
5) The amount of documents for each of the four XML schemas should be checked;
6) Regarding RBD, QA/QC tools should check that at least one ground water body is defined for each river basin district;
7) Regarding referential integrity among the level codes for the different water bodies, QA/QC tools should check the existence and the validity of all codes (e.g EU_CD, DIST_CD, etc).

- *Geometry Harmonisation*

Regarding geometry harmonisation certain QA/QC tools are identified and proposed in terms of topology and statistical accuracy among different layers.
QA/QC tools should:

1) Check the coordinate system for all Member States;
2) Check the X,Y coordinates of each centroid and points/line/polygons of each water body to ensure that they belong to the Member State that they refer to;
3) Check for connectivity and gaps among the main river network and river stretches;
4) Check for existing polygons (polygons closed/not closed);
5) Check if all surface water bodies and protected areas shapes connect and edge-match the country borders and the coastline;
6) Check that adjacent river water body line shapes connect and edge-match within a Member State and across Member State border;
7) Check if river water body line shapes stop/start at the edge of lake water body shapes or if they are continuous;
8) Check for overlaps, slivers, dangles, undershoots, overshoots in and between each water body and protected area.

The Statistical accuracy of the submitted files should also be considered and QA/QC tools should check and provide layer statistics of the accuracy of the lines and the polygons delivered from each shapefile for each water body and protected area of each Member State.
In conclusion data reported through each Member State will be picked up by QA/QC tools that are dealing with data assessment and control. After the successful pass of the QA/QC tools, these data are inserted to the reference database. As a result of the QA/QC tools performed on the reported databases a summary of the main uncertainties and data gaps together with actions identified to address these issues should be indicated and reported over web services back to the responsible Member State.

3. CONCLUSIONS

WISE as a distributed system, in line with INSPIRE and SEIS principles. It is an ambitious and long term project. Currently a process of streamlining is set up integrating a number of different directives and flows to a single system process. A strong partnership is being built with each Member State that is crucial for further development. It should be also noted that several Member States are developing their own water information systems which in some cases are as elaborate and complex as WISE itself (DG Env, 2008).
During the submission of data to WISE a number of practical obstacles were encountered such as incomplete or not submitted data, variable format of the provided datasets and often the data did not comply with the GIS guidance.
The quality of the information submitted by Member States is very diverse and often difficult to read, validate and process. This might be due to the fact that there are often differences between Member States in the interpretation of the GIS guidance as well as the resources they provide for the assimilation of this information. However, a constant improvement of the WISE members is noticed, including a new level of cooperation among DG ENV, EEA, and Eurostat linking the reporting obligations and reporting flows.

REFERENCES

Common Implementation Strategy for the water framework directive (2000/60/EC) . available at: http://circa.europa.eu/Public/irc/env/wfd/library [accessed 5 March 2009].

Cima, J., Feld, H., Haastrup, P., de Jager, A., Jensen, S., Lack, T. and Werner, B., WISE implementation plan [online].European Commission. available at: http://forum.europa.eu.int/Public/irc/env/wfd/library [accessed 5 March 2009].

DG Env, European Commission, Minutes from workshop on WISE. Paris 12/11/2008.

Hannerz, F. and Langaas, S., Establishing a water information system for Europe: constrains from spatial data heterogeneity, *Water and Environment Journal* 21, 200-207, 2007.

Information Systems for Building an ERA in the Field of ICT for Environmental Sustainability

W. Geiger[a], R. Lutz[a], L. Maurer[b], W. Pillmann[c], C. Schmitt[a] , M. Schreiber[d] and K. Tochtermann[b]

[a] *Forschungszentrum Karlsruhe, Institute of Applied Computer Science, P.O.B. 3640, 76021 Karlsruhe, Germany (werner.geiger, richard.lutz, christian.schmitt@iai.fzk.de)*

[b] *Graz University of Technology, Inffeldgasse 21a, 8010 Graz, Austria (lisa.maurer@tugraz.at, klaus.tochtermann@tugraz.at)*

[c] *ISEP International Society for Environmental Protection, Bechardgasse 24/12, 1030 Vienna, Austria (pillmann@isep.at)*

[d] *Leuphana University of Lueneburg, Sustainability Sciences Department & Computing Centre, Scharnhorststr. 1, 21335 Lueneburg, Germany (schreiber@uni.leuphana.de)*

Abstract: To establish a European research area in the field of ICT for environmental sustainability, the ICT-ENSURE project is carried out from May 2008 to April 2010. Within the framework of this project funded by the 7th Research Framework Programme of the European Union (FP7), two web-based information systems are being developed to enhance the exchange of information on research programmes and research results in Europe: an information system on national research programmes in the field of ICT for environmental sustainability and a literature database with publications in this research field, in particular with publications of the EnviroInfo conferences. This paper shall outline the background and objectives of the ICT-ENSURE project. Furthermore, the goals and design of the information system on research programmes will be described and the concept and prototype of the literature database discussed.

Keywords: European Research Area (ERA); ICT for Environmental Sustainability; ICT-ENSURE; Research programmes information system; Environmental informatics literature database.

1. INTRODUCTION

In 2000, the EU decided to create the European Research Area (ERA) (cp. updated strategy in [ERA, 2007]). The European Research Area aims at building a more coherent science and technology environment across the European Union through improved coordination of national scientific and technology capacities of existing and future Member States. For implementation of ERA, the European Commission has launched various ERA support actions within the 6th and 7th Framework Programmes (cp. CORDIS [2009]).

In the 2007-08 ICT (information and communication technologies) work programme, the European Commission has called for proposals for an action to support the building of the European research area in the field of environmental sustainability. A consortium consisting of the University of Technology Graz, the International Society for Environmental Protection (ISEP) Vienna, and Forschungszentrum Karlsruhe (Karlsruhe Research Centre) has submitted a concept for a support action called ICT-ENSURE which has been accepted for funding by the Commission. The support action with a duration of two years started in May 2008. It is based on the activities of the Technical Committee for Environmental Informatics that was founded within the German Society for Informatics (Gesellschaft für Informatik) in 1987 [Pillmann et al., 2006] as well as on the research community in the field of environmental informatics established by this committee in Central Europe.

The ICT-ENSURE concept pursues three main goals (cp. Tochtermann et al. [2008]):

1. Support in fostering a single information space in Europe for the environment (SISE)

   For this, a survey of European research in the field of ICT for environmental sustainability and national research programmes in the EU member states is being compiled. Based on this survey, concepts and recommendations for the creation and further development of an integrated SISE will be developed.

2. Extension of the European environmental sustainability network

   In order to improve information exchange among European experts and programme managers in this field of research, various expert workshops are being organised and the EnviroInfo conferences are being extended to cover additional environmental sustainability topics, with additional EU countries being involved.

3. Development of information systems

   To enhance the availability of information on national research programmes in the field of ICT for environmental sustainability, a corresponding information system is being developed. To improve access to research papers relevant to ICT for environmental sustainability, a literature database is being established. In addition, relevant information will be offered via the website of the ICT-ENSURE project (*www.ict-ensure.eu*).

In the following paragraphs, the information system on research programmes and the literature database of the ICT-ENSURE project will be presented in more detail.

2. ICT-ENSURE INFORMATION SYSTEM ON RESEARCH PROGRAMMES

2.1 Introduction

In the European Union, most of the research programmes in the field of ICT for environmental sustainability are research programmes funded by the European Commission or national research programmes of the member states, including research programmes of regional bodies in the states. The research programmes funded by the European Commission generally are well documented on the Internet. Information can be accessed with the help of the CORDIS information system of the European Commission [CORDIS, 2009]. In contrast to this, information on national and regional research programmes is available in a distributed manner only, if at all. A central information system or single entry point to this information does not exist. Often, the information is available in the respective country's language only. Consequently, it cannot be used by interested persons, who do not speak this language. This results in a risk of double development of research results and prevents them from being used in the best possible way.

The ICT-ENSURE information system is aimed at offering on the Internet meta-information on national research programmes in the field of ICT for environmental sustainability in the member states of the European Union and making this information searchable. Moreover, a single entry point to more detailed information on these programmes shall be provided. Similar information systems for other disciplines were and are being developed under other ERA support actions. An example is the IST research portal CISTRANA. This portal was developed under the CISTRANA project initiated by a European research area working group of member and associated states [CISTRANA 2009].

As outlined in section 1 above, part 1 of ICT-ENSURE is dedicated to the acquisition of meta-information on European research programmes in the field of ICT for environmental sustainability. Experts identify important national research programmes in key domains of the field. National research programmes in these domains are screened in the EU member states, and detailed surveys of the research programmes will be compiled in some selected EU countries. These meta-data represent the basis of the information system.

The meta-information will be opened to the public by a web-based information system. This system comprises a storage component with the meta-data, an authoring component (and author interface) for the distributed acquisition of the meta-data on national research programmes in different countries by the ICT-ENSURE experts, and a presentation component for information retrieval by the end users of the system via the web. The presentation interface will offer various search and navigation facilities, including a navigation feature via a taxonomy of the field, a navigation facility via an EU map for access to the research programmes of selected countries, and a full-text search in the meta-data, including short descriptions of the programmes.

2.2 System requirements

When developing the web-based information system, particular attention has to be paid to an appropriate functionality, a high usability, and a good maintainability of the data. The user, either the user of the public web interface (presentation component), who wants to obtain information about research programmes, or an author of the system, who inserts or updates meta-data of the research programmes, has to accept the way of handling the system offered. Only then are the system developed and its meta-data accessed frequently.

To achieve this goal, several requirements have to be fulfilled:

- The characteristic data of national research programmes in the research field have to be represented. Characterising attributes are the programme name, a short description, country(ies) involved, the institution responsible for the programme, creation, start and end dates as well as ICT and environmental sustainability fields covered (cp. section 2.3). In addition, a research programme may refer to documents in the information system itself (literature documents) or (by web links) to information on the Internet.
- The existing relations and hierarchical structures of the data (relationship of programmes to other programmes, references to legal entities, web links, literature, etc.) must be mapped well to the underlying database system and the user interface to reach a good interaction performance and high acceptance by the end user.
- Several ways to retrieve the desired data have to be provided. For this reason, navigation through countries, research programme hierarchies, identified target fields (keywords) or concerned people will be offered by the system. A comfortable search facility will be provided.
- Different user classes (user roles) have been identified. Scientists, research administration and management staff, promoters as well as any interested Internet user will use the presentation component for data retrieval. Authors of the information system contents and responsible administrators will obtain the role to manage the data via the authoring component.
- Security aspects have to be considered, as the system is open for the public. System access by authors or administrators to perform any data modification has to be restricted by an authentication process

2.3 Software Architecture

Analysis of the requirements led to the identification of several use cases that serve as a basis of the development of the software architecture. Depending on the roles assigned to users (end user, author or administrator), three use cases are distinguished: Restricted access to the data by end users (search and navigate), management of programme meta-data by the author (create, delete, modify, search, and navigate) as well as additional management of system users by an administrator.

Based on the requirements and depending on the use cases, the software architecture follows a common MVC (Model-View-Controller) pattern for a web application. Several feasibility studies have been performed to determine the combination of tools that is suited best for performing the given tasks. The tools selected are shown in Figure 1.

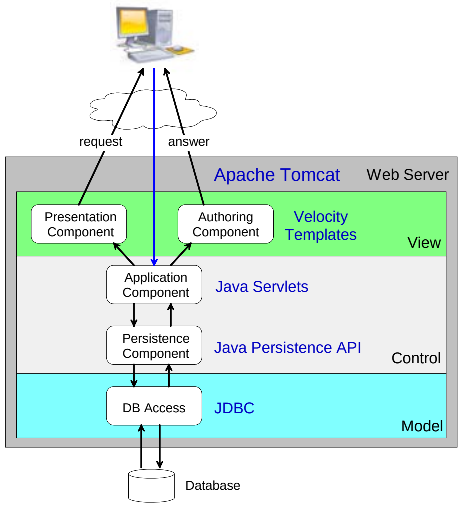


Figure 1. Software architecture and tools used for the information system on research programmes.

2.4 Data Model

Analysis of the use cases and requirements led to the analysis data model shown in Figure 2 in a reduced form. It is modelled with UML (Uniform Modelling Language) to show the dependencies between the classes in an object-oriented way.

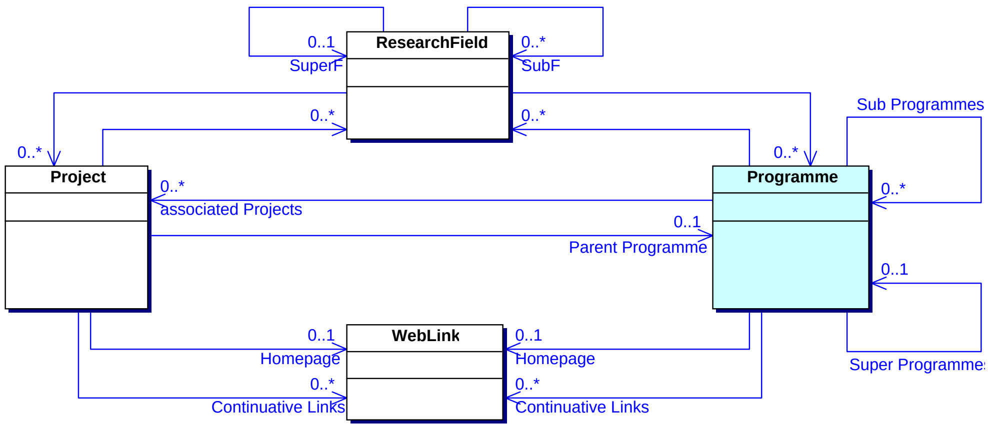


Figure 2. The reduced UML data model for research programmes.

The main class of the data model is *Programme*. A *Programme* may have subordinate *Programmes* and one or no superordinate *Programmes* to which it is attached. Also *Projects* may be included in the information system. A *Project* may belong to a *Programme*, which means that a *Programme* may contain some *Projects* which are performed in its context. Both a *Programme* and a *Project* may have a homepage accessible via a *WebLink.*

The class *ResearchField* serves to characterise the ICT for environmental sustainability field of the research programme. It is referenced by an arbitrary number of *Programmes* and may have superordinate fields and sub-fields to map the hierarchy of the research field.

This class will be taken for navigation through the research fields. In the current implementation, following sub-fields and key values are available:

1. ICT technology area (*Electronics & Microsystems*, *Information Systems & Software, Media & Content*, *Communication Technology & Networks & Distributed Systems* and their sub areas)
2. Environmental sustainability area (*Ecology & Natural Environment*, *Economy* (along the product lifecycle), *Social*, *Environmental Sustainability (general),* and their sub areas)
3. Addressee (*Policy & Justice*, *Public Administration*, *Business & Economic Sector*, *Research & Science*, *Education & Learning*, *General Public*)

For reasons of compatibility with other ERA projects, the areas and sub-areas of the ICT technology area were chosen on the basis of the taxonomy developed in the CISTRANA project [CISTRANA taxonomy, 2008].

Besides the classes above, about 15 other classes were modelled in the design phase to fulfil the requirements specified within the framework of the ICT-ENSURE project.

3. LITERATURE DATABASE

Today, in times of digital libraries and a fast and ubiquitous Internet, scientists may directly access larger data inventories than ever before. However, proceedings of conferences are often contained in libraries in the form of monographies only and, hence, certain contributions to the proceedings and in particular their full texts are not accessible. But this would be desirable, as current developments, research projects, and specialised projects are not only described by journals, but also by these proceedings that may provide scientists with constantly updated information.

That applies to the field of ICT for environmental sustainability too. Large inventories of specialised literature, e.g. the proceedings of the EnviroInfo conferences, cannot be accessed in its full content. Even specialised technical databases, such as the technical environmental library ULIDAT of the German Environmental Authority (*http://doku.uba.de*), do not provide extensive and up-to-date access to the proceedings of the EnviroInfo conferences. The findings are provided with descriptors, but abstracts can be found occasionally only and full texts are not available.

Data mining techniques that are used by the search machine Google for the web and by Google Books in the printing area have resulted in a significant increase in the material available, but so far, conferences and workshops in the field of environmental informatics or ICT for environmental sustainability have not been acquired systematically.

This gap shall be closed largely by the literature database under development. It shall provide the community with access to results in the fields of informatics for environmental protection, sustainable development, and risk management. This field is represented by the corresponding Technical Committee of the German Society for Informatics that also organises the international EnviroInfo conferences. At these conferences that have taken place annually since 1986, numerous scientific papers have been produced:

EnviroInfo conferences	22
Workshops	105
Reviewed papers (conferences and workshops)	> 3.200
Pages documentation	> 30.000
Authors	~ 7.500

It is the paramount objective of the literature database to make the EnviroInfo conference proceedings and the results of working groups of the Technical Committee available online. Single articles shall be available for download true to the original. Access and search will be accomplished via conventional bibliographic data. In addition, the end user can search through the abstract and full texts. Logical search operators of the database will allow for a flexible access to the articles of the proceedings. Hence, far more information will be provided than by a conventional literature database. Access to texts is often aggravated by the fact that the proceedings are sometimes out of stock.

To supply literature in this field to the complete extent - in full text and online - an abstract, the full text, meta-data, and a PDF file of every article have to be generated. As the proceedings and workshop volumes were printed by various publishers, the data are not available in a standardised format. Since 1998, the proceedings have been published in digital form, but sometimes in various formats. Before 1998, no digital versions of the proceedings were published, such that digitisation by scanning with subsequent optical character recognition (OCR) would be required.

The bibliographic data of the results of a literature search may be exported flexibly. For this purpose, the standard export format tab-return or, as a bibliographic format, the exchange format of EndNote or the RIS format may be chosen. The RIS format is on the way of becoming a standard. Many bibliography programmes are supporting it. In the RIS format, the bibliographic data are qualified with tags and can be processed further more easily. Figure 3 shows a screenshot of the prototype literature database.

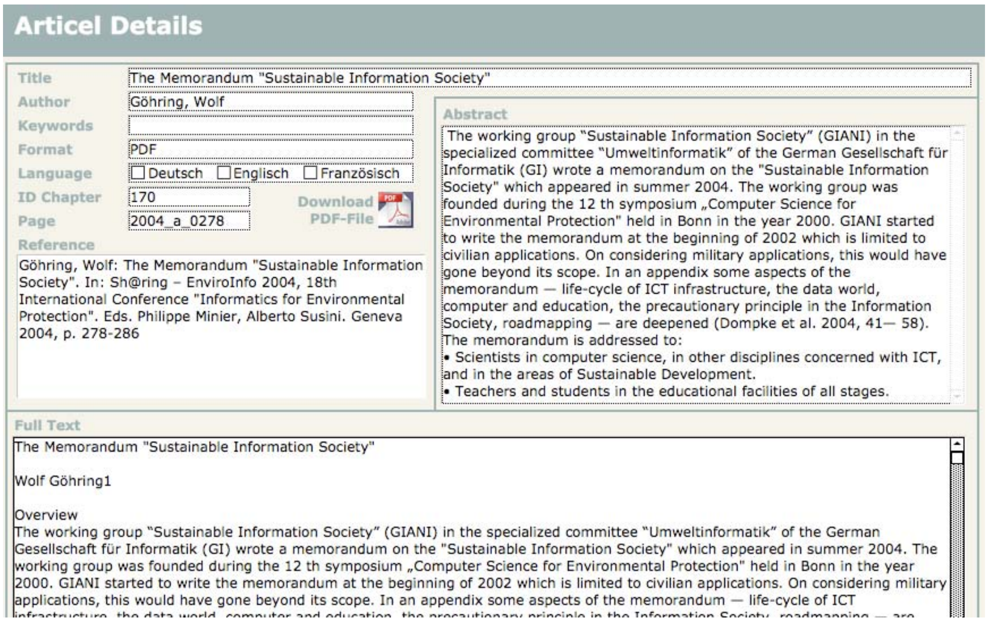


Figure 3. Screenshot of an article with bibliographic data, abstract, full text, and function for downloading the PDF file of the article.

4. STATUS AND OUTLOOK

The two information systems of ICT-ENSURE, which will be made available to the scientific community and research management community in integrated form via the portal of ICT-ENSURE, are currently under development. By the end of April 2009, the prototype of the information system on research programmes will be ready. From September 2009, a first product version will be available to experts to input their meta-data on national research programmes.

A prototype of the literature database has already been implemented. The contributions of the EnviroInfo conference proceedings from 1998 to 2008 are available in full text and with meta-data. By autumn 2009, the literature database shall be implemented on the target system under MySQL, Tomcat, and JSP. The content of 1.250 papers on Environmental Informatics on 10.000 pages shall be included in the literature database.

ACKNOWLEDGEMENTS

The project results have been developed by the ICT-ENSURE project (European ICT Environmental Sustainability Research, *www.ict-ensure.eu*). ICT-ENSURE is financed under the grant agreement number 224017 of the European Commission's seventh framework programme, Theme 3, Information and Communication Technologies.

REFERENCES

CISTRANA, European initiative for the Coordination of IST Research and National Activities, *http://www.cistrana.org/*, 2009.

CISTRANA taxonomy, Revision of the ICT ERA taxonomy, *http://www.cistrana.org/files/cistrana_taxonomy_revision_2.pdf* , 2008.

CORDIS, Community Research and Development Information Service of the European Commission, *http://cordis.europa.eu/home_en.html*, 2009.

ERA, Green Paper on The European Research Area: New Perspectives. Commission of the European Communities, COM (2007) 161, *http://ec.europa.eu/research/era/pdf/era_gp_final_en.pdf*, 2007.

Pillmann, W., Geiger, W., and Voigt, K., Survey of environmental informatics in Europe. Environmental Modelling & Software, 21, 1519-27, 2006.

Tochtermann, K., Granitzer, G., Pillmann, W. and Geiger, W., ICT-ENSURE – A 7th Framework Program Support Action for Building the European Research Area in the Field of ICT for Environmental Sustainability. Möller, A. [Hrsg.], Environmental Informatics and Industrial Ecology,: Proc. of the EnviroInfo 2008, Lüneburg, Sept. 10-12, 2008. Aachen : Shaker Verlag, pp. 456-63, 2008.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Information systems in the International Commission for the Protection of the Danube River (ICPDR)

Alexander Höbart

International Commission for the Protection of the Danube River, Vienna, Austria (Alex.Hoebart@unvienna.org)

Abstract

The ICPDR provides information via the website www.icpdr.org and three additional websites – dedicated to Danube Day, the Joint Danube Survey (JDS) and the Danube Box – to decision makers, journalists, scientists and the general public. The ICPDR additionally provides a Web GIS and several databases. Interested users can self-register to get a password for accessing these expert systems online.

The **ICPDR website** (*www.icpdr.org*) presents an overview as well as in-depth information about the Danube Basin, water-related issues, impacts from different sectors and the work of the ICPDR and related projects. It is continuously updated with information on ICPDR activities, events and related projects as well as with new issues of Danube Watch magazine.

ICPDR databases on the website contain basin-wide datasets on water quality monitoring, point sources of emissions and investment projects.

The **Joint Danube Survey 2** (*www.icpdr.org/jds*), likely the world's largest river expedition, was a 50-day expedition to collect samples of water, sediment, plants, fish and other aquatic life at 124 sites along the Danube and its main tributaries. A fleet of three research ships travelled down the Danube River – from Germany to the Black Sea. The JDS website was launched to accompany the survey with stories, photos and on-board results from the crew. The website now includes the final results, links to many online press articles and some videos about JDS2.

The **Danube Box** (*www.danubebox.org*) is a comprehensive educational tool for teachers and educators working with children of the age between 9 and 12. The didactic material is for use in different disciplines. The website was launched to provide electronic versions of the Danube Box online in several Danube languages, as well as related news and downloads.

Danube Day (*www.danubeday.org*) is a celebrated in all Danube countries each year on 29th June with huge festivals on the banks of the rivers, public meetings and fun, educational events. Over 169 events involving tens of thousands of people took place in 2008. The websites presents an overview of all international events and previous year's highlights. It also links all people involved in the organization of events on national level and in partner organizations to make it one great celebration.

Internal Information Systems of the ICPDR include **Danubis** which provides the delegations and expert groups with an internal working area to share documents, organise meet-

ings and communicate information related to their work. The **Danube Accident Emergency Warning System (AEWS)** is activated whenever there is a risk of transboundary water pollution, or threshold danger levels of hazardous substances are exceed. The system's warning messages by SMS and email are routed automatically to downstream countries and help national authorities put environmental protection and public safety measures into action.

Current focus: Danube River Basin Management Plan (DRBMP)

One of the central elements of the Water Framework Directive is the integrated approach within a river basin. The instrument employed is the River Basin Management Plan, the creation of which depends on the availability of a profound data set that gives information on the current situation of the Danube river. Cooperation beyond regional and national borders thus is an imperative to fulfil reporting and management obligations. Based on these requirements, the **Danube River Basin GIS** (*www.danubegis.org*) provides a platform for exchanging, harmonizing and viewing geo-information and related issues. It is a tool for reporting, management, and planning, while its system architecture remains as flexible as possible to be able to meet future needs.

A dedicated website for informing the public and receiving comments on the draft DRBMP is planned to be launched in May 2009.

Shared Environmental Information System and Single Information Space in Europe for the Environment: Antipodes or Associates?

J. Hřebíček[a] and **W. Pillmann**[b]

[a] *Institute of Biostatistics and Analyses, Masaryk University, 625 00 Brno, Czech Republic (hrebicek@iba.muni.cz)*

[b] *International Society for Environmental Protection, Vienna, Austria (pillmann@isep.at)*

Abstract: The White paper of eEnvironment (Electronic access to environmental information) was elaborated by the Ad hoc Committee on eDemocracy of the Council of Europe (CAHDE) in 2007. Now, eEnvironment is a member in the eFamily like eParticipation and eGovernment and is one of the fundaments of eDemocracy.
In 2008 the Commission proposed a *Shared Environmental Information System (SEIS)* which provides the basis for eEnvironment. The objective is that any citizen can inform about environmental matters and can use this information for active participation in decision making and environmental protection.
During the preparation of EU's 7th Framework Programme, DG Information Society & Media proposed the development of a *Single Information Space in Europe for the Environment (SISE)*. The development of SISE would be essential for the support of the development of implementation tools for the Shared Environmental Information System (SEIS).
Both proposals, SEIS and SISE are aiming at a single integrated information space and a shared environmental information system in which environmental data and information will be combined with knowledge for decision support to foster environmental protection and sustainable development.
In this paper an outline of SEIS and SISE is given and problems and challenges of these two development directions are analyzed. Today the development intensity of SEIS dominates compared to SISE developments. This paper outlines the background of SEIS and SISE, and identifies similarities and disparities in their objectives. The paper stresses clearly that collaboration between SEIS and SISE in the sense of a partnership would be ecologically worthwhile. However, any parallel development, divergence or antagonism would mean a dissipation of valuable resources and a drawback for sustainable development.

Keywords: eEnvironment; SEIS; SISE; GMES; INSPIRE; FP7; environmental informatics.

1. INTRODUCTION

"Information is the currency of democracy". As it is stipulated in Directive 2003/35/EC: "*Effective public participation in the taking of decisions enables the public to express, and the decision-maker to take account of opinions and concerns which may be relevant to those decisions* ...". To achieve effective public participation in the decision-making affecting the environment, the public must have access to environmental information, data and knowledge.

Therefore, the *White Paper of eEnvironment* (Electronic access to environmental information) [Nagy, Legat, Hrebicek, 2007] was presented at the second CAHDE plenary meeting held in Strasbourg on 8 and 9 October 2007. Its fundamental ideas were prepared during discussions of the workshop '*Seamless Access to Environmental Information in the EU - Building an Integrated Information Space for the Environment* 'of the 21st International Conference EnviroInfo 2007 in Warsaw.

The eEnvironment legal basis is the Aarhus Convention, which is implemented in the European Community and supported by the EU Directives: *2003/4/EC Public Access to Environmental Information; 2003/35/EC Public Participation; 2003/98/EC Re-use of Public Sector Information and 2007/2/EC Infrastructure for Spatial Information in the European Community (INSPIRE)* [see also Pick, 2007]. The eEnvironment belongs to eParticipation and eGovernment initiatives of the European Union (EU) and it is predestined to be one of the fundaments of eDemocracy.

The further legal act, '*Towards a Shared Environmental Information System' (SEIS), COM(2008) 46 final,* provides the political, organizational and implementation basis for eEnvironment. It follows from this communication that any citizen has to be informed about environmental matters in order to use this information for an active participation in decision making in the field of environmental protection. The *Directorate-General (DG) Environment (ENV); Eurostat (EUROSTAT),* the *Joint Research Centre (JRC)* and the *European Environment Agency (EEA), the* so-called Group of four (Go4), are controlling the development of SEIS. They are collaborating with the *Global Monitoring for Environment and Security (GMES)* initiative of EU, which represents a concerted effort to bring data and information providers together with users. The intention is that both groups can better understand each other and make environmental and security-related information available to the population who need it by using new enhanced services.

The eEnvironment principle is based on the following ideas [Hrebicek et al. 2008a,b]:

- ***Control effect***: An effective legal protection for citizen implies that decision-relevant information is available for all of them. Access to environmental information provides each individual with the possibility to control the compliance with environmental legislation and to point out deficits in the implementation.
- ***Participation effect***: The right to access environmental information increases transparency and allows a better public participation in governmental decisions.
- ***Education effect*** (awareness effect): Knowledge regarding the state of the environment is not limited to public authorities; hence this leads to an increased public acceptance of measures for environmental protection.
- ***Prevention effect***: The general right of publication of environmental information should discourage potential polluters of the environment, because this bears the risk of negative image effects or legal problems.
- ***Standardization effect***: The European Directive and initiatives dealing with access to environmental and spatial information (INSPIRE and GMES) together with the development of SEIS provide EU-wide comparable principles regarding access to environmental information.

The proposal of a development of a Single Information Space for Europe in the Environment (SISE) was specified by DG INFSO in the Work Programme for ICT research in the 7th European Research Framework Programme (FP7) for 2007/08. Thus, SISE realisation could substantially support the development of implementation tools for SEIS and eEnvironment.

SEIS supported by FP7 research projects developing SISE will provide a kind of an integrated environmental information space in which environmental data and information will be combined with knowledge for decision support of environmental protection and sustainable development.

2. HISTORICAL DEVELOPMENT

2. 1 History of SEIS, the Shared Environmental Information System

Preparatory discussions with EU Member States on possible approaches to improve environmental monitoring and reporting started in 2004 at meetings of the Environment Policy Review Group (EPRG). The Commission presented an assessment of the situation for monitoring and reporting and announced that it was examining how to improve the flow of environmental information, to relieve the burden on Member States and to avoid duplication. The Commission further outlined a vision according to which the information resulting from more coherent monitoring would continue to be managed by the competent authorities in the Member States, but become more readily accessible, including for the public, and shared between all levels of governance, from local to international [Paneli 2008].

In 2005 the Commission outlined a vision for a Shared Environmental Information System (SEIS). It addressed increased sharing and access to environmental information, improvements in monitoring and modernised and streamlined reporting systems.

Activities at EU level to implement this vision have continued jointly led by the Group of four (Go4)[1] in consultation with Member States mainly through the European Environment Agency (EEA) and Eurostat structures. The involved organisations agreed on a Technical Arrangement on the establishment of ten environmental data centres (Öko Institute, 2007).

In order to prepare the Communication, the Commission services presented discussion papers (following in the main the impact assessment structure) containing ideas for actions to implement the SEIS vision at EEA Management Board meetings (2006/07), during the meeting of Directors of Statistical Offices (DIMESA) and in technical meetings with Member State representatives - especially the EEA national focal point (NFP) meetings in 2007.

The implementation work of SEIS started with DG Environment, Eurostat, the Joint Research Centre (JRC) and the EEA, who are primarily interested in and working on the issues that SEIS will address. Subsequently, this group expanded to involve the services in DG Enterprise and Transport (ENTR) leading the Commission's work on GMES.

General objectives of SEIS are:

1. ***Organisation aspect***: Sharing of information with a political commitment, international partnership and networking activities
2. ***Content aspect***: On-line access to environmental information with horizontal (thematic) and vertical (local to global) integration
3. ***Infrastructure and service aspect***: A system with tools for interoperability based on the existing ICT infrastructure, current programmes like INSPIRE, Reportnet, GMES and services for eEnvironment (following Steenmans [2009]).

In February 2008, the Commission adopted a Communication Towards a Shared Environmental Information System (SEIS), COM(2008) 46 final. It foresees that SEIS will be based on a set of principles. According to this Communication, information should be:

- managed as closely as possible to its source;
- collected once, and shared with others;
- available to public authorities;
- readily accessible to end-users to enable them to assess the state of the environment in a timely fashion;
- accessible to enable comparisons at the appropriate geographical scale, and
- fully available to the general public.

Furthermore, information sharing and processing should be supported through common, free open source software tools.

As announced in the Communication, the Commission in collaboration with the Member States and the EEA, is currently preparing a legal proposal focusing more specifically on

[1] DG Environment|, European Environment Agency (EEA), Eurostat and the Joint Research Centre (JRC)

modernising the way in which reporting obligations from environmental legislation are made available using cutting-edge ICT Internet technology.

The proposal will offer a legal basis for an integrated and sustainable EU-wide eReporting System in partnership between the European Institutions (Commission/EEA) and the Member States (MSs). The concept for the EU eReporting system is based on the SEIS principles of a decentralised system and builds on the experience with the implementation of the INSPIRE and the US EPA eReporting System. It will be composed of web-based registers providing access to the information at national level, and of a web-portal operating at EU level. MSs will have the flexibility to build their national registers on a centrally operated content repository or to interconnect existing decentralised information systems and content repositories, held in various locations. The implementation of the eReporting System shall make use of already available infrastructures, both at EU and Member States levels.

MSs and EEA were consulted on an outline of the legal eReporting instrument for the first time in February 2008 at a meeting of the National Focal Points (NFPs) of EEA's environmental information network (EIONET). Since then, the legal proposal has been refined in light of various subsequent consultations, performed by the NFPs, EEA's Management Board, the Director's meeting on Environment Statistics and Accounts (DIMESA) in June 2008, and the SEIS Task Force of MSs established by the Commission services to assist in the overall implementation of the SEIS.

In parallel to this, the EEA has undertaken 21 country visits so far. This has triggered the gathering, sometimes for the first time, of a great variety of stakeholders within individual MSs. This includes public administrations at all levels, state, regional and local, thematic sectors such as water bodies authorities, assessment and enforcement bodies, statistical offices, meteorological institutes and cadastre offices to discuss environmental information management issues in Europe and the general development of eEnvironment services. The legislative proposal based on these consultations is due to be presented by the Commission by June 2009.

In addition, the Prague conference “Towards eEnvironment”, organised under the Czech Presidency in March 2009 presents a memorandum on the SEIS as well as a legal proposal on EU eReporting System which could become the basis for Presidency conclusions on this issue.

Due to its distributed concept, the proposed new EU eReporting System will offer an EU-wide integrated platform where thousands of fragmented environmental information sources can be plugged in. It will substantially improve the quality of available information through data management, as close as possible to the data source.

2.2 History of SISE, the Single Information Space in Europe for the Environment

In 2005 the *i2010 strategy: A European Information Society for growth and employment* was launched by the Commission as a framework for addressing the main challenges and developments in the Information Society and media sectors up to 2010. It promotes an open and competitive digital economy and emphasises ICT as a driver of inclusion and quality of life.

The first objective of i2010 is to establish a *Single European Information Space* offering affordable and secure high-bandwidth communications, rich and diverse content and digital services. Action in this area combines regulatory and other instruments at the Commission’s disposal to create a modern, market-oriented regulatory framework for the digital economy.

The creation of the *Single European Information Space* means addressing the challenges posed by digital convergence. This convergence is to telecommunications what globalisation is to trade - an issue that will affect every domain of governmental activity.

The history of SISE began in 2005 with the formulation of the objective “*ICT-2007.6.3: ICT for Environmental Management and Energy Efficiency*” in EU’s Seventh Framework Programme (FP7). Here, the *Single Information Space in Europe for the Environment*

(SISE) was introduced, in which environmental institutions, service providers and citizens can collaborate or use available information without technical restraints. [Schouppe, 2008]. This will also constitute a platform for eEnvironment.

The aim of SISE bases on an ICT research vision for real-time connectivity between multiple environmental resources, which would allow seamless cross-system search as well as cross-border, multi-scale, multi-disciplinary data acquisition, pooling and sharing. Furthermore, it would allow for service-chaining on the web, thereby stimulating data integration into innovative value-added web services. Targets of SISE were specified during several workshops organised by the Commission in 2007 and 2008.

The last workshop "*Towards a Single Information Space in Europe for the Environment (SISE)*" [2008] was organised by the ICT for Sustainable Growth Unit of DG INFSO and took place in Brussels on 15th February 2008. The goal of the workshop was to provide an outline justification and prioritisation of future research or innovation actions with clear European added value. The overall target was to support an integrated information space and market place for effective environmental management in Europe.

To support the ambitious efforts for integration of environmental information in Europe several projects are now in progress. Examples from the FP7 ICT Work Programme are the GIGAS project (**G**EOSS, **I**NSPIRE and **G**MES **A**ction in **S**upport) and ICT-ENSURE[2], a support action in the area of "ICT for Environmental Sustainability Research". The mission of ICT-ENSURE is to better network national and international environmental research programmes and build up a web based information system for access to the contents of environmental informatics literature and programmes and to ICT relevant sustainability research [Geiger et al., 2009].

3. ELEMENTS FOR SEIS AND SISE

3.1 Elements for SEIS

A key step in implementing SEIS and reaping its benefits will be to modernise the legal provisions on the way, in which information required by environmental legislation is made available. This will be done by revising *the Standardised Reporting Directive* (91/692/EC), which needs to be updated and brought in line with the SEIS principles.

The EEA has a crucial role to play in implementing SEIS. As its mandate is to provide timely and reliable environmental information, it is essential that EEA continues making SEIS the centre of its strategy. EEA is coordinating the implementation of SEIS with the help of its European Environment Information and Observation Network (EIONET).

The EIONET[3] is a network of some 900 experts from over 300 national environment agencies and other bodies dealing with environmental information in 37 European countries; in addition, five European Topic Centres (ETCs) of EEA are working on specific environmental themes. It collaborates with two data centres of JRC and three data centres of Eurostat. Based on input from the EIONET partners, the EEA has identified a set of priority annual data flows. These data in the areas of *air quality, air emissions, inland waters, marine and coastal waters, contaminated soil, nature conservation and land cover*, are used to update the core set of environmental indicators, which form the basis of EEA reports and assessments.

From EEA's perspective, SEIS is covering the following topics:

1. ***Sharing (organisation)***
   - Political commitment (legal framework)
   - Partnership (win-win)
   - Networking (connecting)
2. ***Environmental Information (content)***

---

[2] http://www.ict-ensure.eu
[3] http://www.eionet.europa.eu/

- Horizontal integration (thematic)
- Vertical integration (local to global)
- Online access - real time
- Multi-purpose (policy makers, public, research)

3. ***System (infrastructure and services)***
   - Tools for interoperability
   - Mainly existing ICT Infrastructure
   - Inspire, Reportnet, GMES,...
   - New e-Services for environment

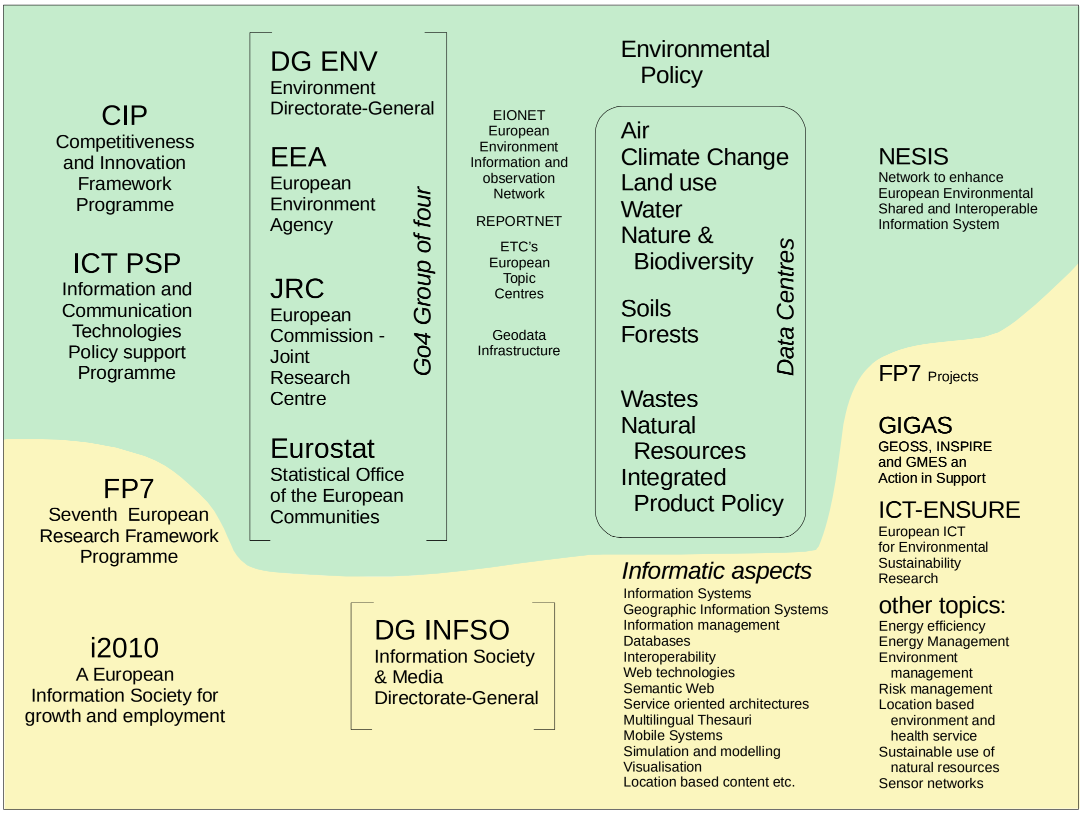


Fig. 1: Programmes, organisations, networks and projects with relevance for eEnvironment

Figure 1 visualizes programmes, organisations, administrative and expert networks and projects with relevance for SEIS (upper part) and SISE (lower part of the figure) on a European level.

Reportnet[4] is EIONET's infrastructure for supporting and improving data and information flows, and has helped to modernise the collection of information and reporting systems [Pillmann, Geiger, Voigt, 2006]. The system integrates different web services and allows for distributed responsibilities. Reportnet has initially been used mainly for reporting environmental data to the EEA, but it is now also hosting some of the Commission's environmental reporting information. The open system also allows for deliveries to be made to other national and international organisations. However, to date the information is often uploaded in the form of reports in word processing or spreadsheet format, which makes it difficult to build databases for efficient retrievals, integration and analysis of the information in order to support policy in decision making.

[4] http://www.eionet.europa.eu/rn

As part of the Reportnet, EEA has created since 1998 a Reporting Obligations Database[5] (ROD) and a Central Data Repository[6] (CDR). In practice it is a pipeline consisting of delivery, quality assessment, aggregation and analysis. Until recently the main streamlining effort was concentrating on the first two segments. The providers know what to deliver, at which time and in which format. The requesters get automatic quality analysis and a feedback channel to the provider. EEA is working on a possible solution towards the aggregation step in the pipeline. It is called the aggregation database and follows SEIS principles in as much as it discovers and harvests datasets on remote systems and aggregates the content in a database while keeping track of the sources. The same mechanism reduces the need for gap-filling of the merged deliveries, like manipulation, because the filling of the gaps is created as just another source that is part of the aggregation, and can be traced [Roug, 2009].

Eurostat has developed the system eDAMIS[7] (electronic Data files Administration and Management Information System) for the SEIS which is the standard solution for transmitting statistical data files to the Eurostat single entry point (see Figure 2).

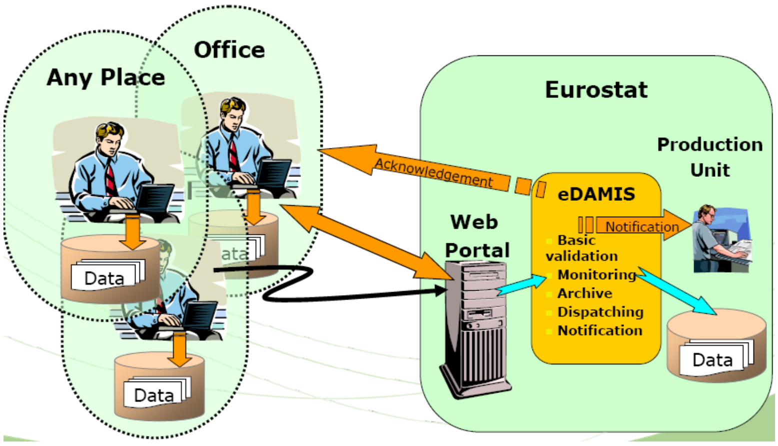


Figure 2: eDAMIS Web Portal transmissions

The portal eDAMIS is an ICT tool which implements the concept of the Single Entry Point (SEP) and simplifies its usage. It provides adapted solutions to several needs and user profiles (NSI - National Statistical Institutes and other organisations), facilitates fully automated data transmissions and guarantees secure data transmission. It offers value added services such as: traffic monitoring, acknowledgements, reminders… No specific infrastructure is needed for using eDAMIS, only an internet browser. The portal eDAMIS (eWA) is currently used by all NSI's of MSs.

Following the Technical Arrangement that has been signed by the Go4 in November 2005 the JRC acts as data centre for soil and forestry (see Figure 1). Therefore, the JRC set up the European Forest Data Centre[8] (EFDAC) which is currently under development.

In the context of EFDAC development, an inventory of forest data holdings in Europe is being carried out. It is focused on the following themes: *Inventory of forest resources, Forest health, Forest fires, Forest products and trade.*

For example, the 'Forest Fires' theme includes forest fire prevention and control; and the use of fire as a cultural tool. There exists the European Forest Fire Information System (EFFIS), the European Fire Database which is focused on the systematic collection of a

[5] http://rod.eionet.europa.eu/index.html
[6] http://cdr.eionet.europa.eu
[7] https://webgate.ec.europa.eu/edamis
[8] http://forest.jrc.ec.europa.eu/efdac/

minimum set of data on each fire occurring, carried out under *Forest Focus Regulation* (EC No 2152/2003) by the MSs participating in the EFFIS network.

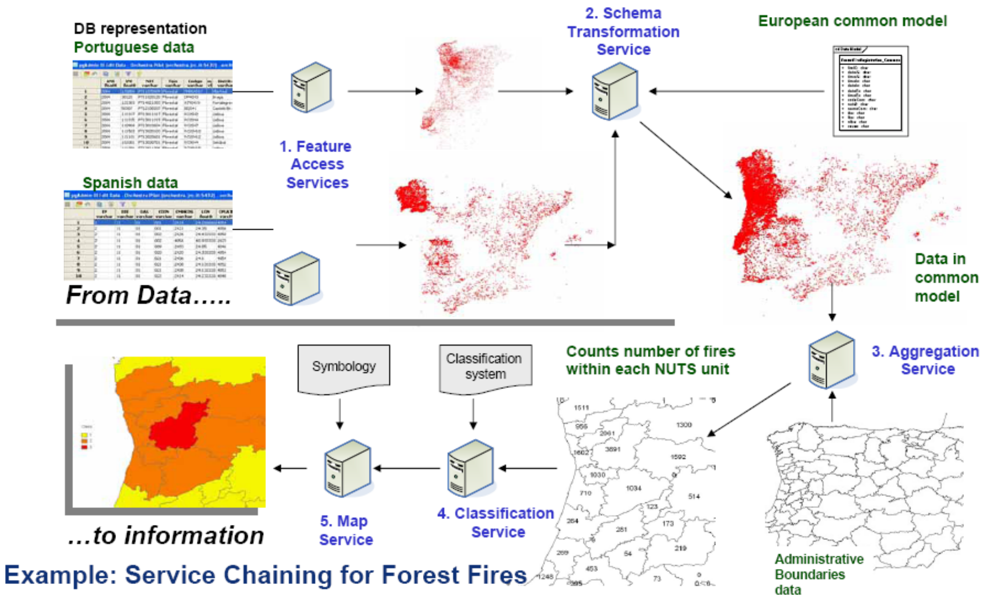


Figure 3: Service chaining in EFFIS

The public access to the database currently allows the users to retrieve general information such as maps of the number of fires for a selected year and for the countries for which data are available (see Figure 3).

3.2 Elements for SISE

SISE is expected to facilitate the access by all actors (decision makers, scientists, users involved in environmental management processes or activities, citizens) to shared qualified information, see Figure 4 [Alerge, Bonnot 2009].

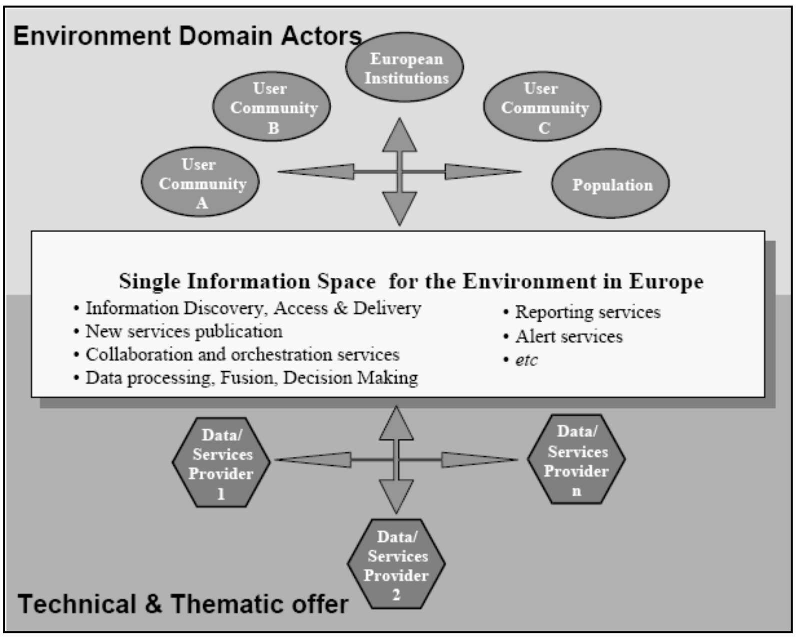


Figure 4: SISE concept [Alerge, Bonnot 2009]

There are recommendations for elements of SISE that came from a DG INFSO workshop. The conclusions drawn for the prioritisation of future work towards a SISE were reported by O'Flaherty [2008]. The following main themes of SISE were identified:

1. ***SISE Application and Services***
   - SISE Services
   - Process Chaining and Uncertainties
   - Real-time Mapping and Modelling
   - Thesauri
   - Open Standards and Open Source Software
2. ***SISE Open Semantics and Standards***
   - Standardisation and Framework Projects
   - Standardisation and Community Knowledge
   - Semantic Web Technologies for SISE
   - Ontologies
3. ***Data Interoperability and Web Communities***
   - Web 2.0 Technologies
   - Data Provision in the Semantic Web
   - SOA/Web Services and Model Driven Communities
   - Social SISE
4. ***Data Visualisation and Modelling including Risk Assessment***
   - Visualisation of Environmental Data
   - SOA and Semantic Web Services
   - Simulation and Modelling
   - Complex 3D/4D models
   - Chained web services and legacy systems
5. ***SEIS Deployment Models***
   - From Framework Projects to Market Deployment
   - Project's Knowledge Loss
   - Regional Application of European Interoperability Standards
   - SISE and Business Models
   - Environmental Information Service Economy

The detailed DG INFSO Unit 'ICT for Sustainable Growth' workshop contributions can be accessed on web[9].

In addition to research activities, the Commission has recognized that ICTs and ICT-based innovations may provide one of the potentially most cost-effective means for MSs to achieve the 2020 target. There are two Communications adopted in May 2008 [COM(2008) 241] and in March 2009 [COM(2009) 111]. They are a first step towards creating a policy framework that will allow proposing in the research of SISE, that the energy-saving potential of ICTs will be widely recognized and exploited. The following prioritisation of topics in the 4th calls of ICT for Sustainable Growth Unit in FP7, CIP[10] and other instruments (such as Calls for Tender) were identified to progress future research and innovation actions on SISE:

Flexible chaining of distributed environmental services

- Methods and protocols for service discovery and chaining
- Automated data fusion, services and quantitative quality assessment

---

[9] http://cordis.europa.eu/fp7/ict/sustainable-growth/workshops/ws-20080215_en.html

[10] http://ec.europa.eu/cip/ The Competitiveness and Innovation Framework Programme (CIP) aims to encourage the competitiveness of European enterprises. With small and medium-sized enterprises (SMEs) as its main target, the program will support innovation activities (including eco-innovation), provide better access to finance and deliver business support services in the regions. It will encourage a better take-up and use of information and communications technologies (ICT) and help to develop the information society. It will also promote the increased use of renewable energies and energy efficiency.

- Semantics, thesauri, ontology services and standardisation
- Chaining of models, predictive tools, including the characterisation of the propagation of uncertainty in chained models
- On-demand distributed geo-processing and services
- Interactive, Web-based 3D/4D analysis and visualisation tools
- Integration of heterogeneous geo-spatial sources, Web 2.0 collaborative community and semantically enhanced services.
- Tools for interactive usable and useful contextualised user interfaces.

This objective is reflected in the work programme for the ICT theme of the FP7 Specific Programme 'Cooperation' defining the priorities for calls for proposals closing in 2009 and 2010 and the criteria for proposals. Research in future and emerging ICT will explore novel scientific foundations to overcome longer-term technology roadblocks and build new synergies between wide ranges of scientific disciplines, like the bases to key future technologies. ICT for environmental sustainability with more emphasis on bridging environmental information spaces and services are supported. Call 4 of the FP7 with its topic "*4.6 Challenge 6: ICT for Mobility, Environmental Sustainability and Energy*" was announced in November 2008. In addition, there is another objective ICT-2009. 6.4 ICT for Environmental Services and Climate Change Adaptation with a special sub-topic *b) Flexible discovery and chaining of distributed environmental services.*

ICT tools for an easy discovery of environmental service nodes on the web and their on demand adaptive chaining (or composition) taking full advantage of international open standards will be essential *elements for SISE*. These will include generic semantics frameworks and dynamic ontology services for the discovery of and the access to distributed environmental resources in a multilingual multi-domain context. SISE will also include methods and protocols for service chaining and for the management of the effects of uncertainty propagation through service chaining.

Generally, projects should be driven by the possibility for a range of users, including non ICT-skilled users, to plug-in their own use cases and get access to customised information and decision support. Solutions shall be validated over different scenarios and allow for continued collaborative development by federated user communities.

The expected impact of projects is a contribution to a SISE in which environmental actors, service providers and citizens can collaborate through improved systems connectivity and semantic interoperability. At the same time, they should contribute to the development of SEIS and support the implementation of the INSPIRE Directive (2007/2/EC).

CONCLUSIONS

The current working plan for the development of SEIS for 2009 to 2010 together with the work programme for the ICT theme of FP 7 *"Cooperation"*, which defines the priorities for calls for proposals closing in 2009/10, show a new synergy approach. SEIS and ICT calls are not antipodes they are tending to join efforts. The objectives for SISE research announced in FP7 with its call "*4.6 Challenge 6: ICT for Mobility, Environmental Sustainability and Energy*" are focused on the support of SEIS development. Therefore, SISE together with SEIS can be considered as advanced associated partners who are on the right track to implement eEnvironment in practice.

ACKNOWLEDGEMENTS

The authors wish to express their gratitude to the European Commission DG INFSO for fostering this contribution with awarding the contract ICT-ENSURE (European ICT for Environmental Sustainability Research), a project under the grant agreement number 224017 of the European Commission's FP7.

ABBREVIATIONS (selected)

CIP	Competitiveness and Innovation Framework Programme (FP)
DG INFSO	Information Society & Media Direction General
DIMESA	Director's meeting on Environment Statistics and Accounts
EEA	European Environment Agency
EFFIS	European Forest Fire Information System
EIONET	Environmental Information and Observation Network
EPRG	Environment Policy Review Group
Eurostat	Statistical Office of the European Communities
FP7	7th European Research Framework Programme
GMES	Global Monitoring for Environment and Security
Go4	Group of four
ICT	Information and Communication Technologies
INSPIRE	EU portal for the Infrastructure for Spatial Information in Europe
JRC	Joint Research Centre, EU scientific and technical research centre
MSs	EU Member States
NSI	National Statistical Institutes
SEIS	Shared Environmental Information System
SEP	Single Entry Point
SISE	Single Environmental Information Space in Europe for the Environment

REFERENCES (Web access 2009/03/10)

Alerge, Ch., Bonnot, S., GENESIS, a flexible solution contributing to the SISE deployment. Proceedings Towards eEnvironment. Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe, (Hrebicek et al. eds.), Prague 25.-27.3.2009. (to appear)

Geiger W., Lutz R., Maurer L., Pillmann W., Schmitt Ch., Schreiber M., Tochtermann K.: Information Systems for Building an ERA in the Field of ICT for Environmental Sustainability. Proceedings Towards eEnvironment. Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe, (Hrebicek et al. eds.), Prague 25.-27.3.2009. (to appear)

Hrebicek J., Legat R., Nagy M., Pillmann W., Challenges of SEIS and SISE: Integrating Environmental Konwledge in Europe. In EnviroInfo 2008. 22. International Conference on Informatics for Environmental Protection. Environmental Informatics and Industry Ecology. Aachen, Germany: Shaker Verlag, 182-189, 2008a.

Hrebicek J., Legat R., Nagy M., Current Trends in eEnvironment and its Role in eDemocracy (Current Trends in eEnvironment and its Role in eDemocracy). In Proceedings of the iEMSs Fourth Biennial Meeting: International Congress on Environmental Modelling and Software (iEMSs 2008). Barcelona, Catalonia: International Environmental Modelling and Software Society (iEMSs), 1612-1619, 2008b.
http://www.iemss.org/iemss2008/uploads/Main/Vol3-iEMSs2008-Proceedings.pdf

Nagy, M., Legat, R., Hrebicek, J., Electronic Access to Environmental Information – an Important Fundament for E-Democracy and Environmental Protection, CAHDE 23 E. 2007.
http://www.coe.int/t/e/integrated_projects/democracy/02_activities/002_e-democracy/

O'Flaherty, J. J., Towards a Single Information Space for the Environment in Europe. Experts Consultation Workshop – ICT for Sustainable Growth – Brussels 15 Feb 2008.
ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/sustainable-growth/sise-workshop-report-08_en.pdf

Öko-Institut, St. Zangl, J. Lohse, H. Stahl, D. Schüler, C.O. Gensch: Implementation of Environmental Data Centres. Öko-Institut 2007. http://www.oeko.de/oekodoc/477/2007-081-en.pdf

Pick Th.: From Århus to Inspire: Putting Environmental Information on the Map. Proceedings EnviroInfo 2007, Warszawa, O. Hryniewicz, J. Studziński, M. Romaniuk (eds.), Shaker Verlag Aachen 2007

Pillmann, W., Geiger, W. and Voigt K., Survey of Environmental Informatics in Europe, *Environmental Modelling & Software*, 21(11), 1519-1527, 2006.

Paneli M., Good Decisions Need the Right Information at the Right Time: The Shared Environmental Information System for Europe (SEIS). EnviroInfo 2008. 22. International Conference on Informatics for Environmental Protection. Environmental Informatics and Industry Ecology. Lüneburg. 2008. http://www.enviroinfo2008.org/programme_overview.php

Roug, S., An Introduction to the Aggregation Database, Proceedings Towards eEnvironment. Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe, (Hrebicek et al. eds.), Prague 25.-27.3.2009. (to appear)

Schouppe, M., Towards a Single Information Space for the Environment in Europe. Concept and purpose. Expert Consultation Workshop, European Commission Information Society & Media Directorate-General, Unit ICT for Sustainable Growth. Brussels, 15. February 2008. ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/sustainable-growth/workshops/workshop-20080215-mschouppe_en.pdf

Steenmans, Ch., The Future of Environmental Information. Power point presentation 28.-29.1.2009, Wallingford. http://www.ceh.ac.uk/SEIS/documents/EEASEISCEH.pdf

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

The Finnish Environmental Information Portal

M. Hynninen[a], **R.Teiniranta**[a] **and Y. Sucksdorff**[a]

[a] *Finnish Environment Institute*

Abstract: Finnish Environment Institute (SYKE) acts as the national centre for environmental data in Finland. The data stored in SYKE's information systems is widely used for environmental monitoring, environmental modeling, forecasting and impact analysis.

The Data and Information Centre of SYKE gives services for the whole environmental administration in Finland (some 25 locations with 2500 employees) and has the responsibility for

- General ICT management
- Development and management of information systems
- technical support and training on information technology
- harmonization and development of SYKE's international reporting

The environmental information collected in SYKE covers a large amount of nationwide environmental data and is managed by three different sub-systems created for internal use.

1) Environmental Information System (HERTTA): Continuous collection, management and reporting of environmental register databases (updated time series). Data and information includes e.g. monitoring of water quantity and quality, environmental protection, biological diversity, land use and environmental loading.

2) Environmental GI-System: Centralized Citrix® ArcGIS-environment for distributed spatial data production and services. The system has about 500 users and contains both SYKEs datasets as well as data from other providers covering the whole of Finland. Spatial data is stored in Microsoft® SQL Server databases with ESRI ArcSDE.

3) Remote Sensing: Satellite image processor and systems to deliver data to users in different formats. Operational remote sensing products include snow covered area, lake ice snow coverage area, sea surface temperature, surface algal blooms, turbidity, chlorophyll-a and land cover data. Most of the products are near-real-time products.

All the above sub-systems produce data that is accessible for general public from the Internet portal Oiva, opened in May 2008 (www.ymparisto.fi/Oiva). Through this Portal access to datasets owned by the Environmental Administration and which have no legal restrictions is free for all registered users. New data is transferred daily to the portal from operative systems, sensitive information in data is filtered out before publishing. Portal includes data viewing, reporting and download functions.

Future

Changes in the data policy together with the common interest for environment, new technology and the need for e-infrastructure have created new demands and new possibilities to upgrade the current portal. The aim is to facilitate the use of the large amount of environmental data which SYKE is holding for different purposes like research, education, reporting and spreading environmental information.

SYKE is continuously improving the possibilities to use SYKE's data. There are several tasks ongoing at SYKE to facilitate the use of our data, of which the most urgent are:

• Ensuring easy and open access to register and spatial data (data warehouse development)
• Planning and building of INSPIRE view and download services
• Upgrading of spatial datasets and metadata

When upgrading the existing environmental datasets and information systems, national and international standards, regulations, guidelines etc. are taken into account. New standards and data specifications ease the harmonization of data and the implementation of new services.

lenses are the major sources of water supply in many atoll islands in the Pacific and Indian oceans, particularly in dry seasons. Presently, numerous two- and three-dimensional mathematical models are available for the simulation of atoll island aquifers. The two-dimensional models such as the powerful SUTRA are unable to represent the three dimensional distribution of various parameters, the boundary conditions of the problem, and adequately simulate pumping wells. These limitations may be overcome by using a three-dimensional model, however, this will be a very challenging task. To demonstrate the related problems, an attempt was made to simulate the case of Home Island in the Indian Ocean. This exercise demonstrated that such modelling required a very fine three-dimensional discretisation of the island and short time steps of a few hours, in order to overcome the numerical instability. This required a very significantly large CPU time, even on the most sophisticated workstations. Obviously, this problem can be overcome by running the model on a super computer. However, the main problem is due to the paucity of knowledge of various parameters involved. While the Home Island dataset is comprehensive, the quality and quantity of the available data proved inadequate to meet calibration needs.

Keywords: Portal, Spatial data, registers, remote sensing

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Framework Conception for the Environmental Information System of Baden-Württemberg (Germany)

A. Keitel[a], R. Mayer-Föll[b] and A. Schultze[c]

[a] *Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg, Postfach 10 01 63, 76231 Karlsruhe, Germany, andree.keitel@lubw.bwl.de*

[b] *Umweltministerium Baden-Württemberg, Postfach 10 34 39 Stuttgart, Germany, roland.mayer-foell@um.bwl.de*

[c] *Datenzentrale Baden-Württemberg, Postfach 30 03 22, 70443 Stuttgart, Germany, a.schultze@dzbw.de*

Abstract: The Environmental Information System of the federal state of Baden-Württemberg in Southwest Germany is the organisational, technical and content-related framework for the processing of environmental information within the state and municipal administration in Baden-Württemberg. It also provides the basis for reporting to the public. The main emphasis of the Framework Conception 2006 for the Environmental Information System is to describe the structures and processes which were developed due to a major reform of the administrational structures in Baden-Württemberg – as a joined solution for a comprehensive UIS operated in partnership by state and municipal administrations.

The most important factors for the success of the Environmental Information System of Baden-Württemberg (UIS BW) are outlined with concrete examples of organisational and technical solutions.

Keywords: Environmental Information System.

1. GOAL OF THE FRAMEWORK CONCEPTION AND ITS AREA OF APPLICATION

The Environmental Information System of the state of Baden-Württemberg in Southwest Germany is the organisational, technical and content-related framework for the processing of environmental information within the state and municipal administration in Baden-Württemberg. It is part of the e-Government Concept Baden-Württemberg as well as of the ICT cooperation between the state and the municipalities on the one hand and the state, the Federation and the European Union on the other hand.

The Framework Conception 2006 (RK UIS 06) describes the functional, legal, political and organisational factors – effective at the time of its development – which determines the operation of the Environmental Information System (UIS BW). The RK UIS 06 shows the potentials of the application of modern information technology and makes recommendations for the further development of the UIS BW.

A special characteristic of the UIS BW is its cross-departmental scope and area of application. An appropriate project organisation guarantees the constant participation of all involved departments through the Ministry of the Environment, which is in charge of the UIS BW. Like the preceding Framework Conceptions of the years 1986, 1990 and 1998, the RK UIS 06 has been approved by the government of Baden-Württemberg.

Due to a reform of administration in 1995, which caused the devolvement of environmental tasks from the state to the municipalities – especially in the field of water management and waste management – the communities and the rural districts already took part in the development of the Framework Conception of 1998. At the same time, the municipalities were committed by public law agreement to the implementation guidelines of the conception.

At the beginning of 2005 a law came into force, which caused a major reform of the administrational structures in Baden-Württemberg, and an even greater transfer of responsibilities from the state authorities to the administrative authorities on the regional and the district level. As a result it was necessary to rearrange the collection, management and exchange of environmental data within the UIS BW and to provide appropriate ICT services, suited best under these circumstances. The RK UIS 06 is significantly characterised by the new demands of the ICT cooperation between the state and the municipalities.

The data management and communication within the UIS BW also has to ensure, that environmental information can be passed on from the UIS BW to other federal states and the federal government, especially in order to fulfil the obligation of reporting to the European Union.

In addition to the collection and provision of data for the law enforcement in the environmental authorities of the state and the required reporting to national and supra-national authorities, there is an obligation to inform the public comprehensively about the environment. The Environmental Information Act of Baden-Württemberg, which implements the demands of the directive 2003/4/EC of the European Parliament and of the Council on public access to environmental information on the state level, places further demands on the UIS BW. These also had to be taken into account by the RK UIS 06.

2. CONTENTS OF THE FRAMEWORK CONCEPTION

The chapter Inventory (Bestandsaufnahme) shows the important information systems and services of the UIS BW. About 75 different well matched hardware and software components are described in three categories: base systems, specialised applications and comprehensive components. The registry of services comprises of 18 services, which are either used by several specialised systems or made available as web services.

The chapter Technical Design (Technisches Konzept) describes the service architecture of the UIS BW in detail. Further emphasis is put on the processing of spatial data as well as on common technical standards and safety issues.

The chapter Information Management (Informationswesen) deals not only with technical and organisational aspects of the management of data, documents, information (especially metadata), knowledge and business processes, but also with workflow management and data quality assurance.

A separate chapter concerns with the diverse cooperations (Kooperationen) between the state and the municipalities, between the federal states and the federal government and between the public administration, research institutes and companies. All these cooperations are essential for an economical development of the UIS BW.

Two further chapters, Data Privacy and Economy (Datenschutz, Wirtschaftlichkeit), focus on the special demands concerning data privacy and profitability of the UIS BW, with a view to e-Government.

Last but not least, the chapter Implementation (Umsetzung) of the RK UIS 06 makes recommendations on how the strategic goals can be transformed into concrete solutions, which benefit the users of the UIS BW on all levels.

3. MANAGEMENT AND EXCHANGE OF ENVIRONMENTAL DATA WITHIN THE ICT COOPERATIVE BETWEEN THE STATE AND THE MUNICIPALITIES

The uniform gathering and management of environmental data, derived from the execution of environmental tasks, is regulated within two major ICT projects, in which the responsible ministries, the urban and rural districts and the governmental districts take part. The Information System for Water, Immission control, Soil, Waste, and Occupational safety and health (WIBAS) – under the responsibility of the Ministry of the Environment – and the Nature Conservation Information System (NAIS) – under the responsibility of the Ministry of Food and Rural Area – were introduced into usage for all concerned authorities of the state and the municipalities.

WIBAS and NAIS have a joint software architecture and data organisation. They comprise about 35 different specialised applications and 12 services, used by both systems.

The specialised applications and the services are developed and maintained by the Centre for Information Technology of the State Institute for Environment, Measurements and Nature Conservation (LUBW) and the Data Centre Baden-Württemberg (DZBW) in a coordinated approach. Usually the software is updated once a year. Together with the specialised data and the spatial data from the central database of the LUBW, the software is delivered to all institutions in a unique setup. The on-site support for the local authorities is carried out by the DVV BW, a group of institutes, providing ICT services especially to the municipalities. The Information Technology Centre for the State Administration of Baden-Württemberg undertakes the support of the governmental district authorities.

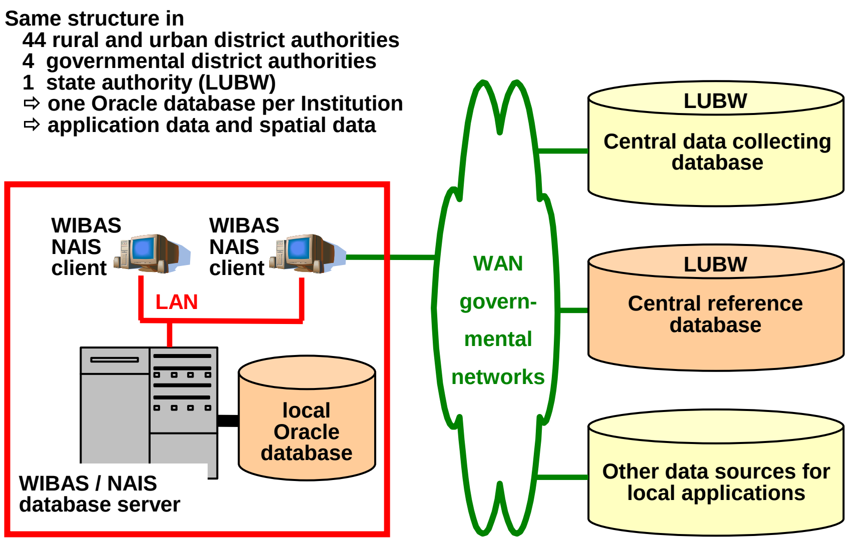


Figure 1. WIBAS / NAIS: System architecture and data access

The WIBAS Data Catalogue (WIBAS-OK) provides the policy for the data management regarding the content and organisational measures. It defines in detail, which data have to be collected by whom, how the collection should be carried out, and where the data should be sent.

The data collection and the primary data management take place mostly in the local databases of the authorities. The schema of the Oracle database is distributed in combination with the yearly WIBAS software update (Fig. 1).

Once a month the data from the local databases, which is destined for the data exchange between the administrations, is transferred to the central UIS reference database, located in the LUBW. The technical and organisational aspects of the data exchange are precisely stipulated in the WIBAS-OK. All users of the institutes can access the reference database read-only with the UIS Reporting System, according to their rights appointed in the WIBAS-OK. Currently 3,200 users are able to access the reporting system by use of the administrative networks of the state and the municipalities. The UIS-BRS is based on the Software CADENZA, which is developed in cooperation with the disy Informationssysteme GmbH, a company located in Karlsruhe, Germany. In Germany, CADENZA is used by six federal states, three larger administrations of the federal government and partly in the municipalities.

Like WIBAS the processing of spatial data within the UIS BW includes central and local components (Fig. 2). The Centre for Information Technology of the LUBW merges spatial base data from the land survey offices with specialised spatial data from the LUBW and other sources in a central data pool (RIPS-Pool). Oracle Locator is used for the data management, ArcGIS/ArcSDE for the central organisation of spatial data and the provision of appropriate services.

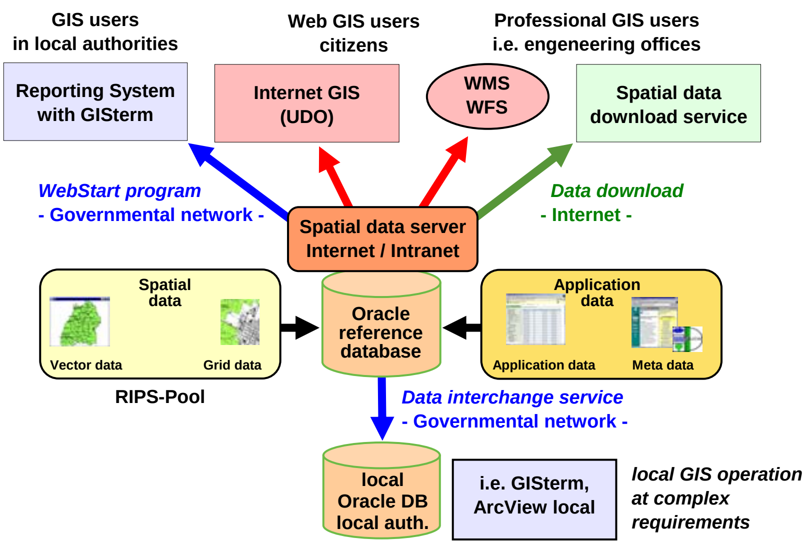


Figure 2. Management and distribution of spatial data

The environmental spatial data of the RIPS-Pools are provided by the LUBW and other authorities with a state-wide responsibility as well as from the monthly data exchange of the enforcing authorities. The data is disseminated from the central spatial data server to the different user groups in the following ways:

- to users of GIS (geographic information systems) in the offices through the UIS Reporting System or GISterm and GISterm Web, components based on the CADENZA software,
- to citizens with their own GIS infrastructure by means of web services (WMS, WFS) or the download service for ESRI shapefiles,
- to the common citizens with help of the web application Environmental Databases and Maps online (UDO), which is also based on the CADENZA software.

RIPS makes a significant contribution to the Spatial Data Infrastructure of Baden-Württemberg (GDI-BW), which is currently developed and serves the implementation of the directive of the European Parliament and of the Council establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) on the state level.

4. CENTRAL MANAGEMENT AND DEPLOYMENT OF DATA BY THE LUBW

The LUBW not only runs the UIS reference database and the RIPS-Pool, as a central data pool for spatial data, but also the central database for measuring series – the Measuring Series Operation System (MEROS). MEROS, which was put into operation several years ago, provides a uniform modelling for all measuring data of the state-wide monitoring networks operated by the LUBW, covering various areas like water, soil, air and radioactivity.

A unique characteristic of the UIS BW is the fact, that the three major data bases – the UIS reference database, the MEROS measuring series database and the RIPS-Pool – have been combined by ER modelling to one logical database, the Database for the Comprehensive UIS Components (Fig. 3). All application data and spatial data of this encompassing reporting data base can be accessed by the UIS Reporting System.

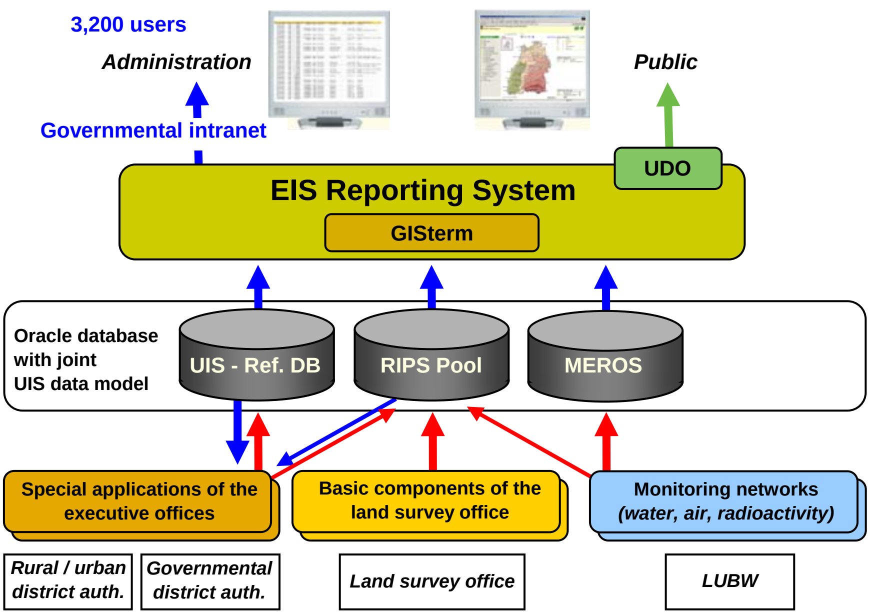


Figure 3. Central databases and Reporting System of the UIS BW

According to the status and the special needs of the users working with the UIS Reporting System, three different versions of the system have been deployed:

- A full version for specialists in the environmental administration, who need specific data analysis.
- A less complex web version (BRSWeb) for occasional users in the administration without special knowledge, who want to access and visualise data quickly over the administrative networks.
- A particular version of the BRSWeb, called Environmental Databases and Maps online (UDO), which offers public access to a limited amount of application data and spatial data over the internet. This version especially helps to fulfil the demands of the Environmental Information Act of Baden-Württemberg.

5. INFORMATION PORTALS – SERVING THE ENVIRONMENT AND THE CITIZENS

The legal basis for the reporting of environmental issues to the public is the Environmental Information Act of Baden-Württemberg. The obligation to provide environmental information demands a high level of quality and transparency of the data presentation. It is very important, that the information is comprehensible not only for the specialists. Considering the amount of data, features like a clear structure, intuitional user guidance and convenient search functions have a great significance.

The means of choice in this case is the implementation of internet portals, which enable transparent presentation with a combination of reports, documents, catalogue information like metadata, databases and news.

The prime example for an environmental information portal is the German Environmental Information Portal (PortalU) which is developed and operated by a cooperation between the federal government and the federal states.

PortalU regularly collects all environmental information of the state authorities in Baden-Württemberg available on the internet and provides the public with different options to arrange and search these data.

A main characteristic of PortalU is that the public data provided by the various authorities all over Germany are described in a uniform manner in the Environmental Data Catalogue. The data descriptions (metadata) from the UIS BW naturally can be found in this catalogue.

At present it is not the intention of PortalU to provide data from the municipalities. This gap is filled by the Environmental Information Portal of Baden-Württemberg (Portal Umwelt-BW), in which all providers of environmental information on the state level as well as the larger cities in Baden-Württemberg take part with their specific websites. Stringent regulations prevent inconsistencies and duplication of work concerning PortalU. Therefore all information providers on the state level are held in the same way both in Portal Umwelt-BW and PortalU. The structuring, the metadata and the tagging of PortalU have been adopted by Portal Umwelt-BW.

Portal Umwelt-BW also helps the municipalities to fulfil their obligations regarding the Environmental Information Act of Baden-Württemberg in the best way possible. In the meantime this concept – including the administration tool and the search engine (Google Search Appliance) – has been adopted by the responsible ministries of the Federal States of Saxony-Anhalt and Thuringia. A corresponding cooperation between these two ministries and the Ministry of the Environment of Baden-Württemberg was started.

6. WEB SERVICES

Environmental information is disseminated by the use of web services at an increasing rate both to users within the administration and to public users over the internet. The underlying technique allows transmitting data – and partly functions as well – without redundancy and always up-to-date. But it is necessary to set up the services in accordance with the standards of the World Wide Web Consortium (W3C) for web services in general and the Open Geospatial Consortium (OGC) regarding web map services.

Today the LUBW already provides a selection of about 60 standardised web map services (WMS). Anyone running a web server, therefore, can integrate corresponding up-to-date maps from the UIS BW dealing with subjects such as nature conservation, water management and climate protection into their own website.

Due to the increasing amount of web services, provided by the UIS BW to environmental offices in Baden-Württemberg, an automated registry had to be established. Consequently, the LUBW has developed such a registry in conformity with the UDDI standard, and especially for the reporting services. The web map services, which are at the same time developed for the Spatial Data Infrastructure of Baden-Württemberg (GDI-BW) according to the INSPIRE standards, are also registered in a special Catalogue of Metadata and Services (RIPS-OK), which conforms to ISO 19115.

One service of the LUBW, frequently used by environmental offices of the municipalities, concerns the reporting services of the web application Environmental Databases and Maps online (UDO). This service enables the environmental administrations of the district authorities to integrate reports, which are generated using the corresponding data of the UIS reference database (see chapter 3) within the area of the district, into the websites of the municipal information providers (Fig. 4).

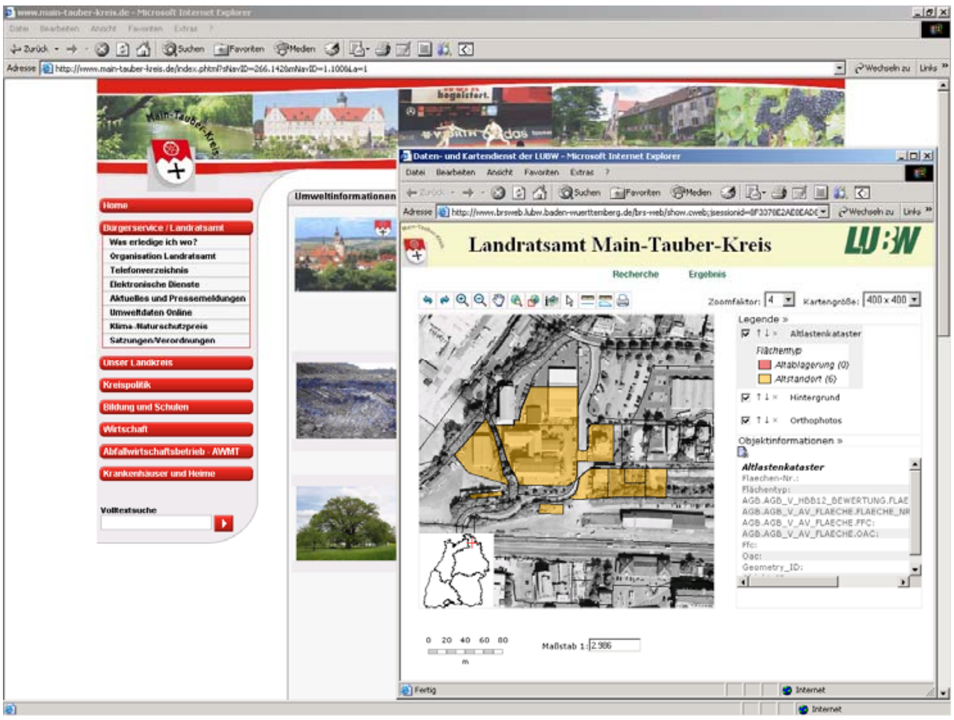


Figure 4. Thematic map from UDO in a web presentation of a rural district authority

7. PERSPECTIVES

The Framework Conception for the Environmental Information System Baden-Württemberg (RK UIS) contains guidelines and recommendations, which shall either be implemented gradually over the next years or taken in to account when components of the UIS BW need to be developed or modified. A major update of the conception is conducted normally every five to ten years.

The demand for an update is primarily owed to a modification of the determining factors, such as new laws or further changes in the administrational structure, which causes a shifting of responsibilities.

The technological progress and the market trend in ICT open up new possibilities, which in the first instance have to be considered in a timely manner as well as in pragmatic way, but in the medium term need to be evaluated in a strategic and conceptual manner.

The RK UIS 06 is based on a long-term experience of developing and operating an environmental information system in an intensively structured federal administration with distinctive cooperative elements. Therefore, it may be of use for institutions and authorities of member states of the European Union with a comparable allocation of responsibilities.

ABBREVIATIONS

LUBW	State Institute for Environment, Measurements and Nature Conservation (Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg)
MEROS	Measuring Series Operation System (Messreihen-Operationssystem)
NAIS	Nature Conservation Information System (Naturschutz-Informationssystem)
PortalU	German Environmental Information Portal (Umweltportal Deutschland)
RIPS	Spatial Information and Planning System (Räumliches Informations- und Planungssystem)
RK UIS 06	Framework Conception for the UIS BW of the year 2006 (UIS-Rahmen-konzeption 2006)
UDK	Environmental Data Catalogue (Umweltdatenkatalog)
UDO	Environmental Databases and Maps online (Umwelt-Datenbanken und -Karten online)
UIS BW	Environmental Information System of Baden-Württemberg (Umweltinforma-tionssystem Baden-Württemberg)
WIBAS	Information System for Water, Immission control, Soil, Waste, and Occupational safety and health (Informationssystem Wasser, Immissionsschutz, Boden, Abfall, Arbeitsschutz)

REFERENCES

Rahmenkonzeption 2006 des Umweltinformationssystems Baden-Württemberg: http://www.uis.baden-wuerttemberg.de/servlet/is/29835/

Konzeption RIPS 2006: http://www.uis.baden-wuerttemberg.de/servlet/is/29847/

Konzeption WIBAS 2006: http://www.uis.baden-wuerttemberg.de/servlet/is/29852/

Kooperative Entwicklung wirtschaftlicher Anwendungen für Umwelt, Verkehr und benachbarte Bereiche in neuen Verwaltungsstrukturen (F+E-Vorhaben KEWA) – Forschungsbericht Phase III (2008): http://www.uis.baden-wuerttemberg.de/servlet/is/41180/

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Process-Oriented Modeling and Infrastructure for Remedial Decision Support System

Jan Pavlovič[a]

[a] *Faculty of Informatics Masaryk University, Brno, Botanická 68a, Czech Republic (pavlovic@fi.muni.cz)*

Abstract: This work introduces a new concept of enterprise architecture, which combines the principles of Service-Oriented Architecture and Business Process Management Suites. This concept is focusing rather on the process as a basic artifact than on the service. It is called Process-Oriented Infrastructure. Process-oriented infrastructure concept was applied in the project Waste Site Energy Management Calculator, which is a mutual project with Masaryk University and U.S. EPA. It resulted in a business process-oriented application, which is capable of dealing with remedial processes that are computably demanding. The infrastructure enables the run of Genetic Algorithms computing the parameters for the remedial processes.

Keywords: energy management, business process management, process-oriented infrastructure, remedial technology

1. INTRODUCTION

This work is motivated by mutual cooperation with Faculty of Informatics, Masaryk University and U.S. Environment Protection Agency (EPA). The agency came with a task concerning a decision support system for remedial technologies.

During the analyses of the specified problematic I have identified several key issues that need to be dealt with.

At first it is necessary to theoretically define the remedial strategy process. There is a huge number of parameters entering the remedial process. That is why it is essential to find an appropriate technique from the Artificial Intelligence domain, which will be capable of handling the problematic.

The designed system needs to be able to use Software as a Service principle and has to be compatible with the existing ICT software environment. This demand requests the adoption of Service-Oriented Architecture (SOA). And it is also necessary to define a suitable subset of the Service-Oriented Architecture since it is a very complex solution

The system has to cope with Business Process Management problematic to satisfy the remedial process management requirements.

The solution will lie in the Enterprise Architecture domain. At the same time a new process-oriented infrastructure and a conformable modeling technique which will satisfy the needs of this problematic has to be designed.

2. PROCESS-ORIENTED MODELING AND ARCHITECTURE

This modeling technique is designed to help organizations with adopting the Business Process Management (BPM) principle. Basic building block is represented by business process. It can be called process centric or process oriented. Adoption of this technique is continuous, which mitigates the risk of failure and enables continuous improvement and sustainability.

The main goal is identification of the strategic business processes. Although there is a complex theoretical background in the BPM area, it is still difficult and expensive to fully adopt BPM principles. Migration of every possible business process to the Business Process Management System (BPMS) would be very expensive and risky. That is why there will be a focus on the strategic business processes only. The rest of the processes will not be implemented within BPMS. This will help with avoiding common problems similar to the SOA problematic.

I came to this conclusion during my practices in the BPM area. Especially converting process from the designed state to the executable state in Business Process Execution Language (BPEL) is complicated and does not bring the desired results, which should be simplification and extensibility for future modification. The maturity of BPEL engines is not ready for fast development yet. It is difficult to gain orientation in the graphical representation of BPEL where business process is complicated as well. There are the main reasons which create the necessity to find balance between putting all processes to the BPMS and hard wiring them in the programming languages.

It is possible to incorporate other business process candidates it the next iterations when it is necessary.

After the first step – identification of strategic business areas it is necessary to identify candidates for business processes. These candidates are specified and modeled in the Business Process Modeling Notation (BPMN) and the essential management metadata are added.

Only those business process candidates are chosen which conform to the criteria specified in the candidate test step. Problematic candidates need to be avoided and not implemented into the BPMS rather then risking the failure of process functionality.

2. 1 POMA Life Cycle Flow

Overview model is described on Figure 1. It represents the overall sequence of steps that specify, which processes will be implemented in the BPMS and which not.

In the SOA world, the economics of funding determines what processes get funded to be developed. This is why it is needed a set of criteria by which can be assessed, which candidate processes are a priority for implementation.

Not all candidate processes should be implemented as process; there may not be enough budget for that. But the functionality of the processes that fail to pass the criteria test is still needed. Therefore those processes still need to be implemented as part of some other component or realized by an application.

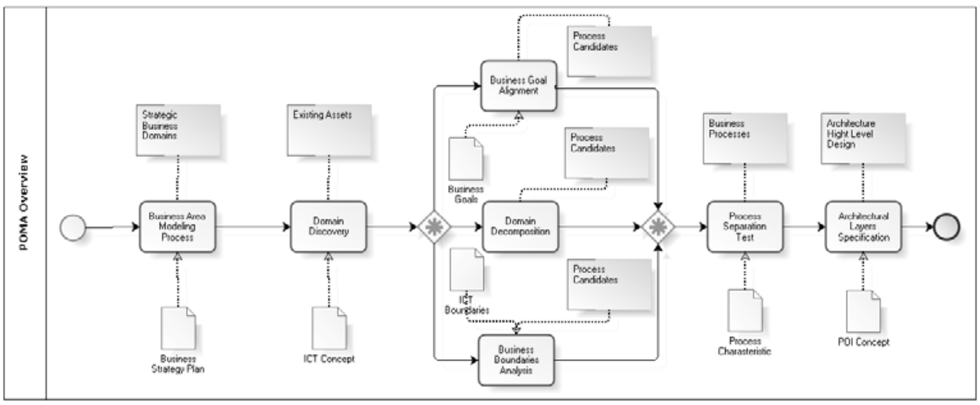


Figure 1: POMA overview process schema

3. CASE STUDY

Nowadays, efficient information and communication technology (ICT) brings us possibilities of using complex environmental mathematical models to find the best solution of these problems. I started the collaboration with U.S. EPA in 2003 to develop the energy management decision support tool. One of the results of this mutual collaboration is on-line solving ICT tool: Waste Site Energy Management Calculator Pavlovič – Mahutová (2006). With this tool you can compare several remedial technologies according to input constraints. Each remedial technology is represented as an energy equation developed by U.S. EPA 3RMA (2003).

There was a pilot version of the Calculator developed at the beginning of the cooperation with U.S. EPA. The basic implementation of remedial principles was tested on this version.

3.1 Problem Domain Definition

This case study is a interdisciplinary project. Nowadays computer science concepts as well as concepts from environmental science are used here.

3.2 Remedial Problematic

The polluted territories with highly toxic pollutants belong into nowadays painful ecological problems. Since they have a direct impact to the human health it is important to remove the pollution and recover the original state. Remedial process is highly expensive and energetically demanding as well. Currently there are lot of the remedial technologies. Each of these technologies is highly specific to the exact contaminant or the group of the contaminants and their efficiency is bind with the type of the medium. Several technologies have to combined to remove the entire contamination of the territory. But the amount of these combinations is vast. The theoretical solution for the optimal combination of remedial technologies with predefined constraints is problematic and there exist no satisfying solution.

Hazardous waste cleanup sites create health hazards by the nature of their contamination. In addition, they adversely impact human health by air emissions generated during site characterization and remedial processes. Treatment of polluted sites is complicated, expensive and energy consuming process. Waste cleanup sites are significant environmental threats and the process of cleaning them up should be carefully planned so that no additional environmental problems are created. There are numerous energy options within the waste cleanup sector such as the reduction of energy consumption during characterization and remedial process, the reduction of green house gas emissions, the avoidance of secondary impact on human health, the enhancement of waste minimization and energy recovery 3RMA (2003).

4. REMEDIAL PROCESS

Ecological risk assessment and its impact on human health is the essential functionality for the next steps that lead to the remediation of pollutants. However, it represents only a small part in the whole system. Figure 2 is describing the whole remedial process cycle. The EPA 3MRA system considers the modeling problem from the source of contamination to the risk assessment.

On the basis of risk assessment and its impact on human health it is necessary to process counteractions and remove stresses from the location. A particular problem domain will be modeled first. Thereby possible remediation steps are discovered that are mainly dependent on the medium where the pollutants are located & their concentration and their variety. This model is mathematically described. The remedial strategy is formed using above mentioned data and the priorities of the remedial process defined by domain experts. This strategy offers several possible variations of how to decontaminate particular location. The follow-up 3MRA procedure is performed after the strategy is applied.

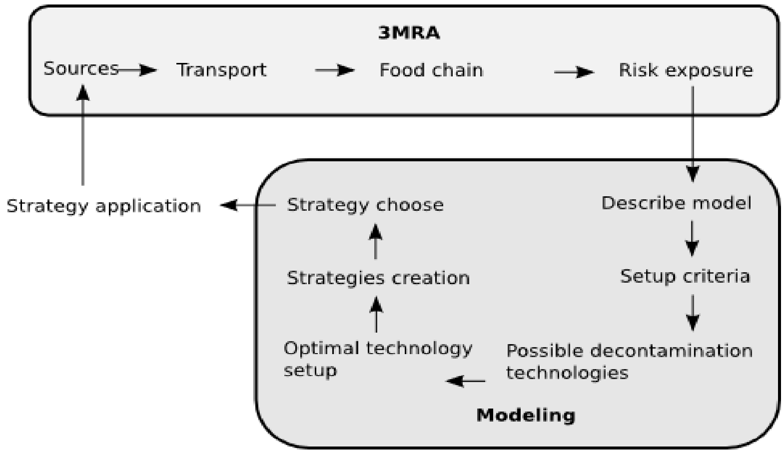


Figure 2: Remedial process cycle

5. CONCLUSION

It is recognized that cleanup actions at federal and State Superfund, RCRA Corrective Action, Brownfields, and voluntary cleanup sites often have significant energy requirements over many years. Very few sites include efficient energy management in the original design and operation plans in spite of the fact that sites offer a significant potential to save and/or produce energy Figure 3.

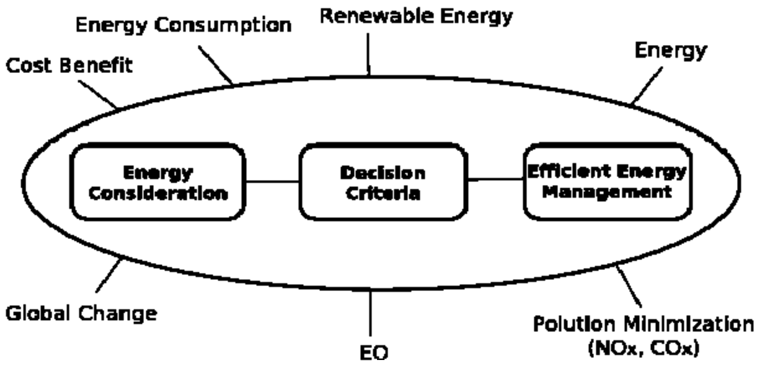


Figure 3: Calculator benefits

In response to the lack of assessment tools to determine the energy requirements, pollution and greenhouse gas potentials at waste cleanup sites and to evaluate energy efficient remedial technologies, Decision Support Tool for Remedial Energy Management is a tool to assist to site managers in assessing the energy implications of remedial actions proposed or completed at waste sites.

The tool is a good high-quality web-based software tool, which can be used for decision making, site management and educational purposes in relation to promote to energy efficiency at waste sites.

REFERENCES

3RMA 2003. U. S. EPA Center for Exposure Assessment Modeling. The Multimedia, Multi-pathway, Multi-receptor Exposure and Risk Assessment

Pavlovič, Jan – Mahutová, Katarina. 2006. Energy Management at Waste Clean up Sites, Avoiding Secondary Air Impacts to Human Health. In *Environmental Health in Central and Eastern Europe*. U.S.A. : Springer, 2006. 270p. ISBN 1-4020-4844-0.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

SISE impact research in The Netherlands, the Flevoland case of Legal Mapping

R. Peters[a], **T. van Engers**[b] **and B.Woudenberg**[c]

[a] *University of Amsterdam and Feed project (R.m.Peters@uva.nl)*

[b] *University of Amsterdam, Leibniz centre for law (T.M.vanEngers@uva.nl)*

[c] *Zenc (Barry@zenc.nl)*

Abstract: For many reasons citizens, businesses and civil servants need access to regulations. The traditional approaches to provide access to these regulations are not satisfactory to these users, who have to cope with vast amounts of often interfering regulations. Normal questions like "where will I be able to do this kind of activity" or "will this activity be allowed here" are hard to answer in traditional web-based services environments. There are many attempts to create one-stop-shop front-ends to eGovernment, but these are seldom built from the perspective of the user. The Legal Atlas offers an innovative approach that has already shown it can be beneficial. It resembles Participatory GIS (PPGIS) but it is claimed here that Legal atlas is a powerful addition to these kinds of transparency tools. This paper describes a number of prototypes that have led to the set-up of the real life application. The Legal Atlas approach has been designed to follow semantic web standards and object oriented legislative drafting to enable negotiations and conflict resolution in the Dutch Province of Flevoland. The discussion about Legal maps has become more relevant in the light of the INSPIRE directive and the authors contributed to the INSPIRE impact analyses of the Dutch regional authorities. The claim of the research is that INSPIRE based map layers may be counter productive to conflict resolution because of the tendency to be too general or too specific with no dynamic adjustment possibilities based on flexible regulative parameters. By linking maps to their legal base and providing easy access to the application area's of the regulations the relevant actors may see more of each others negotiation space. The alternative system described here is designed to support INSPIRE environmental policy implementation, where qualified map layers can help to avoid conflict polarization.

Keywords: legal maps; conflict resolution;

1. INTRODUCTION

In this paper we will show an eGovernment approach that helps citizens, businesses and civil servants to cope with a vast number of rules in their day to day decision making processes. The approach is based upon standardized legal knowledge representations and maps. We advocate that the approach explained in this paper offers a powerful means to ensure client satisfaction, especially when there is a strong interaction between the building blocks of European laws and local regulations. A number of prototype systems have been built earlier in the Addwijzer project and subsequent designs to prove our proposed solution conceptually. The first prototypes contained state-of-the-art navigation techniques that helped to navigate through four layers of laws and regulations: European, national, regional and municipal. Early results from using and demonstrating the system in the Netherlands showed not only high user satisfaction but also gave us the confidence that this type of solution would be feasible in other Europe's countries. Later prototypes were made in collaboration with Dutch provinces, especially with the province of Flevoland around issues concerning the environment. The prototype building was accompanied with a broader scope of all Dutch provinces having to deal with the INSPIRE directive. The researchers involved were responsible for the Dutch regional impact analyses investigation for INSPIRE. This scope provides some insights in technological and political problems when dealing with SISE developments in Europe.

2. METHODOLOGY

The research follows an iterative design tradition from software development rather than an empirical survey approach. A number of prototypes, screen mock-ups and real life systems have been tested by users from government and businesses. Most testing groups were between 10 and 20 individuals and different interviewing techniques have been applied to obtain thorough feed back.

The first highly structured tests have been carried out by a team from Ireland lead by dr. David Newman from QUB. The second structured range of tests has been carried out by the university of Amsterdam. Both sets have been recorded on videotape, transcripted and analyzed. Additionally a number of questionnaires have been used.

The second range of tests followed the Software usability measurement inventory (SUMI) approach. We do think this approach is more catered for usability testing rather than for affordance testing in the line of Gibson's school of functionality design, however. The last set of structured interviews will focus on affordance needs in the INSPIRE production chain. Between the structured interviews over 30 interviews have taken place with experts from Dutch regional authorities, large project investors, Inspire experts , GIS developers, NGO's and city managers. The prototypes have been subject to over 10 presentations for legal practitioners and eGovernment consultants. Most of these interviews and discussions have helped to fine tune the functionality for the next prototype.

2.1 PPGIS

The current extent of Public Participatory Geographical Information Systems (PPGIS) programs in the US has been evaluated by Sawicki and Peterman. They catalog over 60 PPGIS programs that aid in "public participation in community decision making by providing local-area data to community groups" in the United States (Craig et al, 2002:24).

PPGIS systems usually provide one detailed area or issue for a certain period of time in which interested actors can discuss the issues with each other. Examples of such issues are building a new square in town or the realization of a subway trajectory. More elaborate systems deal, for example, with permanent balancing of the different interests of stakeholders living in a region with high historical value.

The Legal Atlas Approach differs from previous PPGIS approaches in four ways:

- Comprehensiveness; entire regions can be represented for eParticipation purposes on a continuous bases;
- Citizen orientation rather than government orientation: the legal domain at hand, which consists of both existing legal constraints as well as policies under discussion, is modelled and so finding opportunities and threats are supported.
- The semantic RDF architecture is maintainable and interoperable in such a way that domain and cross-language cooperation is made relatively easy.
- Conflict resolution : the user applies map techniques to discover area's of negotiation and conflicting legal constraints

The improved access to, and therefore comprehensiveness of, the legal domain at hand is perhaps the most important distinctive factor which allows that our solution, contrary to stand alone and domain issue-based PPGIS, can be part of the normal primary process and the continuous decision making that is going on in all town halls and regional authority offices.

2.2 Legal Atlas

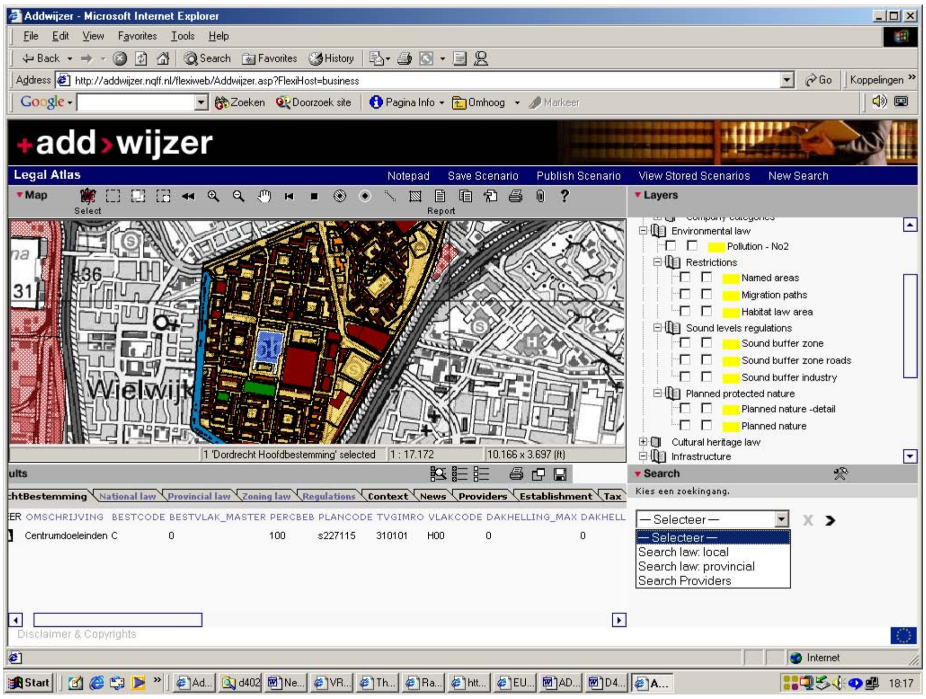


Figure 1: result set showing city centre purposes area in a neighbourhood in Dordrecht based on IMRO classification

The Legal Atlas[1]© approach aims to enable clients to search in local and regional regulations in a way that is comparable to querying databases. The systems should enable the ultimate transparency of local regulations in the sense that the citizen could ask “where can I do this?”, instead of: “can I do this in place such and such?”. The where-can-I-do-this query requires much more than a good website from a well-organized municipality; it requires some way of standardizing local regulations across all municipalities of a region or country. More specifically it requires that we can aggregate the legal information of all those municipalities and the region into one system. This system should then allow citizen to find answers to case related questions during the democratic decision making processes. Such an eParticipation application consequently requires a unique coding for all local and regional legal functionalities affecting citizens and a mechanism for semantically mapping between the legal terms and legal procedures affecting those required functionalities of citizens. In that way the client of the service could view his case in the light of aggregated data and monitor the consequences of decisions and opposition. This brings a new perspective on maintainability and consistency of the ever-growing legal system as a whole.

A second version of the application was designed using the experiences from the first version. In the second version we wanted to extend the retrieval possibilities of law from text to map, in other words; by clicking on a certain regulation the user would be able to identify all objects or map area’s where this regulation applies by means of a visualisation mechanism. While this may seem a small extension of the first version, the reader should realise that there is a big difference between the question: “what will be allowed here?” and

[1] The legal atlas was developed in the context of the ADDWIJZER project of the eContent programme

"where will this be allowed?". The second inverse question demands a complete semantic definition of the world and complete legal coverage of that world as previously described.

2.3 User consultation and Evaluation of Legal Atlas©

This prototype, which was developed in an eParticipation project called FEED, makes use of IMRO in combination with the GEMET[2] "GEneral Multilingual Environmental Thesaurus". It presents 5.298 descriptors, including 109 Top Terms, and 1.264 synonyms in English. The 5.524 terms belonging to the parental thesauri and not included in GEMET, constitute an accessory alphabetical list of free terms. GEMET is increasingly relevant because of INSPIRE related democratic decision making processes all over Europe. The regional layers of data and most of the content are provided by The Province of Flevoland [3] who maintain a number of databases (Thematic Spatial Data Infrastructure) about the economy, the environment, our cultural heritage and the public infrastructure. This Regional information is aggregated content by nature. For crossing different regions and

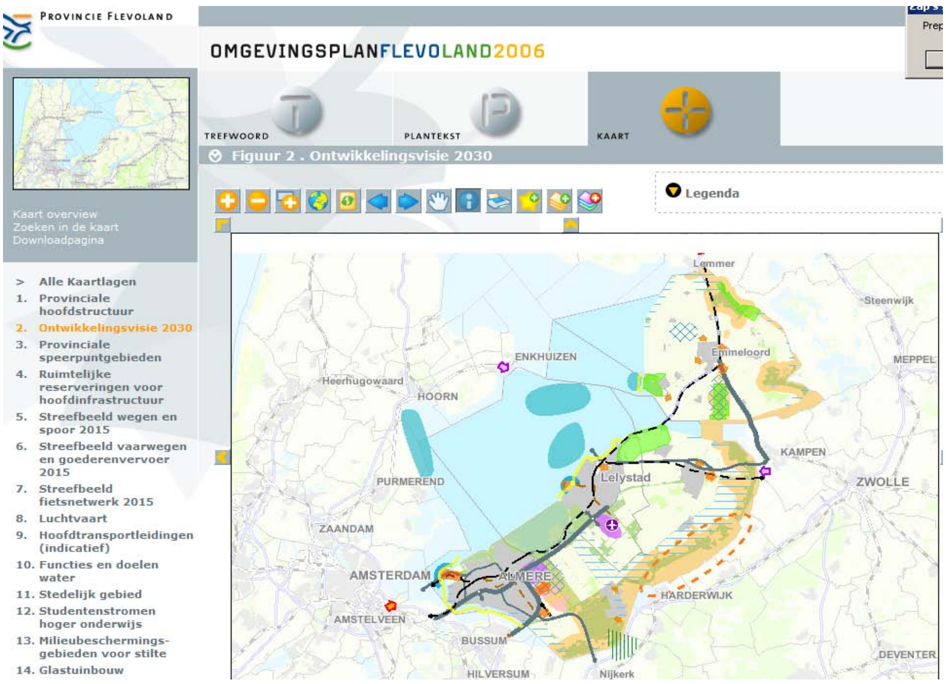


Figure 2: The Flevoland real life participation system based on Legal Atlas approach

countries we apply GEMET[4].

The coloured area's are text-to-map area's that can be activated using keywords out of the controlled vocabulary. The next step is to concentrate the retrieval possibility on a much more narrow scale, see Fig. 3.

There were a number of reasons to choose the Flevoland area for the testing of the Legal Atlas Approach;

The Flevoland public authority has to coordinate the revising of the entire regional development plan with six cities in a concurrent fashion. All changes are made available to

---

[2] *http://www.eea.europa.eu/*

[3] with thanks to Mary Haselager, Kees Kersten and Jolanka Perk of Geopolis Flevoland

[4] This was advised by Prof. Hrebicek of Masaryk University, Czech Republic, member of the Feed team

all players by webservice. Flevoland launched the "Geopolis" initiative to create the right spirit of collaboration among actors involved in informing the public and businesses about redevelopment plans. Another reason is that Flevoland has had the vision to write the explanatory paragraphs on the 200 pages regional development plan in XML and in object oriented slices. This means that businesses and citizens or NGO's can find relevant and precise location based interpretations of the usually very diplomatic language of such plans without having to rely on professionals who are paid to read through political documents.

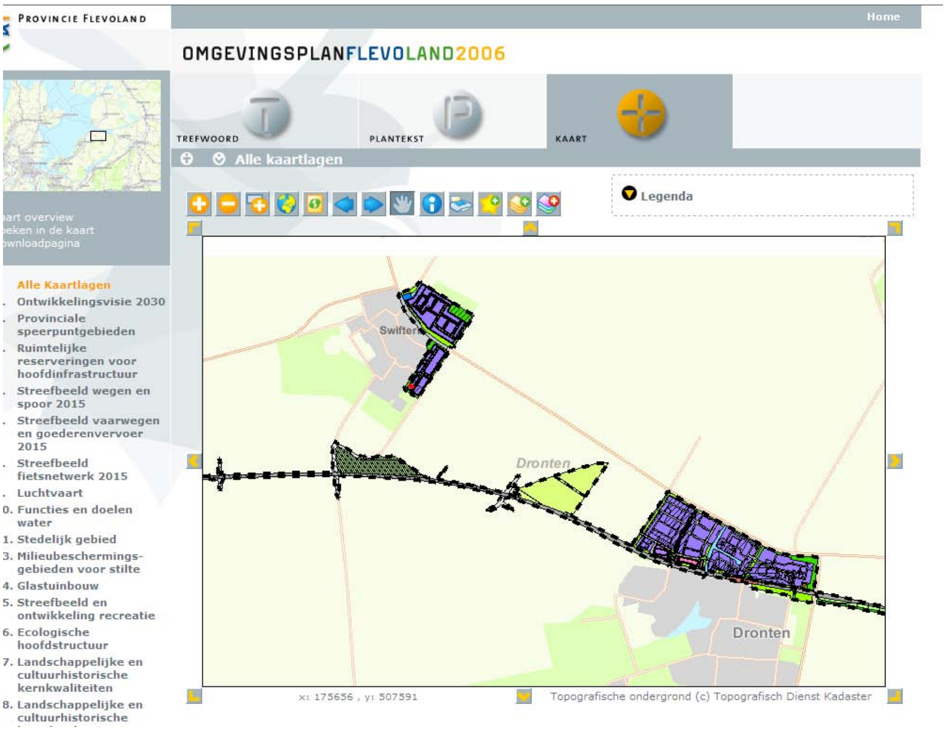


Figure 3: detailed text-to-map retrieval

The third reason is that the area to be re-developed is under high pressure from many actors who want to launch economic and recreational activities in a densely populated area where one European directive demands that people stay away to save protected birds, while the other European directive demands large scale redesign operations to maintain water quality in the inner lake. Flevoland has become the focus of the European Feed research project.

The primary goal of the FEED project is to test the legal/semantic approach in a real life environment and to test the assumptions of earlier prototypes in the core business arena of democratic production.

The second goal is to further investigate the remarkable difference in orientation between public authorities on the one hand, and citizens and businesses on the other, that address a gap that cannot be easily bridged by technology. The third goal is to test the GEMET semantic network to enable conflict resolution methods where "on the fly" translation of search words, folksonomies and decision phases is required. This part will be tested as part of the eParticipation FEED[5] project. Examples of cases where such questions are relevant are NIMBY problems or environmental trade-offs. The NIMBY problems foreseen in the Flevoland use case scenario's are issues like the allocation of a water ski resorts near parts of the ecological protected surroundings, hotel and other recreation facilities and the construction of new islands for housing. The legal reasoning becomes relevant when the law offers limited possibilities of intrusion of "green" area's by compensation of other "green"

[5] http://www.feed-project.eu/

elsewhere. For the purpose of European transferability it is intended to apply Google maps technology rather than GIS technology to support "contour editing" by visitors of the website. The visitor can indicate a specific area by drawing a line on the map to discuss and to support his case with documents, research reports, legislation and plans that can be uploaded , stored and downloaded for that drawn area .

3. SISE PORTAL INTERFACE DESIGN

While building the earlier prototypes of Legal Atlas it was made clear that contextual information of the conflict are is very well appreciated. The information part of the ADDWIJZER interface has been improved on a number of aspects:

- -Upload of resources are made possible using a map area;
- -Text-to-map area queries are enabled;
- keyword lists are maintained using controlled vocabularies in protégé, enabling fnding all area's ith certain legal characteristic
- a query for contextual documents can be applied for a certain map area
- on the fly conversion from ESRI-Oracle to Google is made possible to maintain relation between front- end service delivery and back office maintenance;

The design of the interface resulted in the figure below:

Figure 4: interface of Legal atlas with additional retrieval for contextual information[1]

The interface of the prototype shown above is linked to the Oracle database of the regional authority Flevoland. All map layers available can be projected on this Google map screen. This way the conflict resolution system will maintain dynamic input from legal constraints and map layer updates. The document retrieved in this particular screen dump example shows the relevant research report as "hit" upon entering the keyword algemeen_rec in the search field. It is also restricted to the bleu square area on the map. Resources can be uploaded and downloaded this way. An important part of the interface is the geographical interface with Google earth and Google maps. The connection between the geographical data and the metadata that describes them can be seen in figure 4. The Map has a number of different regions which are of some type of land use. These regions have particular geographical representations in the form of layers with polygons.

3.1 Supporting Conflict situations

Maps are regarded as a mean so support politics. Visualisation of data should inform participants in decision making about facts and it should support administrations with enforcement, permit processing and policy adjustments.

The Dutch Regional Authorities have explicitly stated that Geographic visualisation of environmental constraints should support policy making and democratic decision making:

"*The supply of Provincial spatial data should be designed in such a way that it contributes substantially to the political and administrative role of the Provinces and its administration by the year 2011 and that it especially supports the citizen and businesses in their geographic information needs*[6]". INSPIRE can be regarded as one of the key stones of the Dutch regional efforts in this ambition, but there are some issues with transparency beyond the obvious arguments against all kind of transparency. The "vision" part of the portal is about solving conflict-situations. Figure 5 describes the architecture behind the solution we implemented to resolve the conflict. A description of the plan in the middle top box is combined with the normative descriptions in the regulations. Both the regulations and the plans are about SKOS vocabulary and Regions, and it is possible to trigger particular situations that are allowed or disallowed. It is important to note that the OWL models do not become inconsistent, but that the allowed and disallowed situations are also OWL classes. It is now possible to represent the disallowed situations and its regions as conflicts on the map.

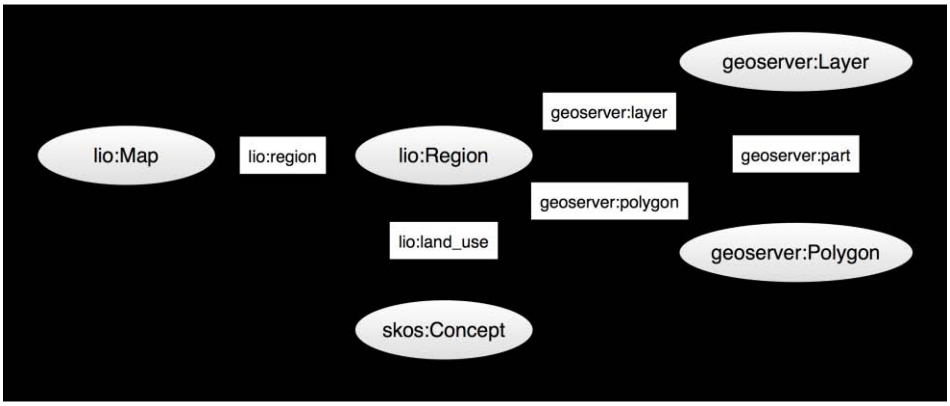


Figure 5a: The interaction between the maps, its regions with particular land use and the geographical representation.

[6] proGideon project 2009

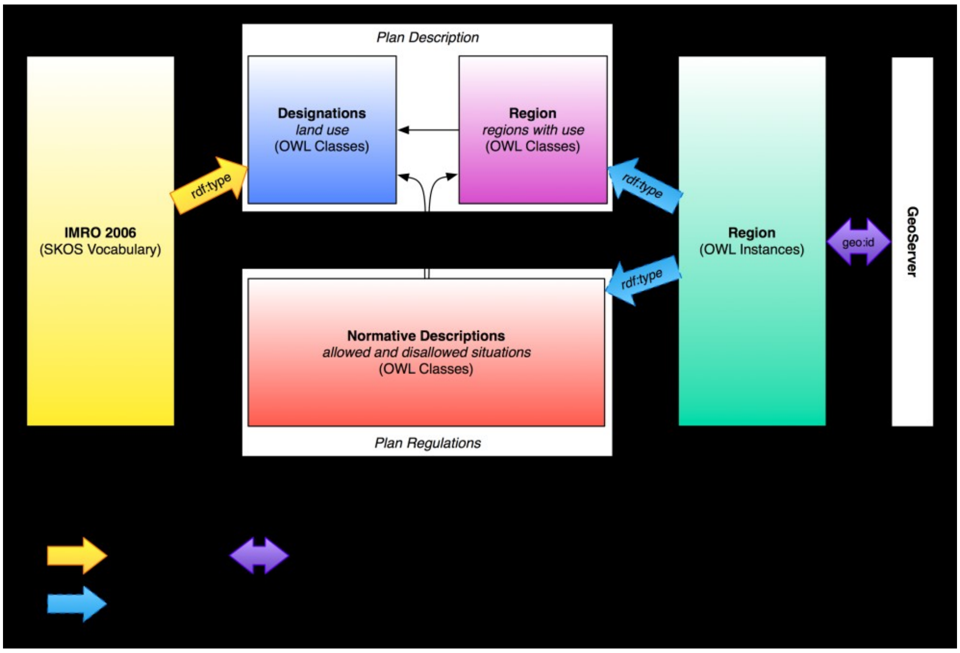


Figure 5b: The interaction between the maps, its regions with particular land use and the geographical representation.

4. PROBLEMS WITH INSPIRE MAPPING FOR DUTCH REGIONAL AUTHORITIES

The Dutch Provinces have carried out an impact analyses of INSPIRE[7]. This impact analyses mentions several ambiguities with the new infrastructure, Starting with the definition of "Protected sites" according to the International Union for the Conservation of Nature (IUCN):

A Protected Site is an area of land and/or sea especially dedicated to the protection and maintenance of biological diversity, and of natural and associated cultural resources, and managed through legal or other effective means'.

The broad scope of this definition does raise a discussion among policy advisors how to apply the INSPIRE directive wisely. If too many "sites" are defined it tends to weaken enforcement possibilities and reporting to Europe and SEIS may become tedious. If too few area's are being identified as protected the Regional authority will loose those "sites" it wants to protect. The latter occurs because of the hierarchical effect of European directives. "Lesser" legal regimes become less useful to oppose against economic interests. The Ecological main infrastructure (EHS) in the Netherlands, that was meant to become the vehicle to ensure less fragmentation of green area's in an urban country, may well become negligible in it's effect given the greater impact of Natura2000 area's.

Some Map layer standardisation efforts turn out to be disastrous for regional authorities. The Dutch provinces have several examples of non-intended effects. The regulation of salt and sweat water balance, for example, indicated that one Province was obliged to subsidise farmers for lack of sweat water while those farmers had one of the best crops in Europe.

Sudden environmental Transparency can even have financial repercussions as effect. Showing health risks depicted on a map may lower the price of houses and the owner might enter a claim against the province for bad public management. Another range of difficulties regarding transparency derive from bad design. The Dutch have decided to avoid having Ecological main structure data available at high zooming levels. It suggests a granularity with legal consequences that was never intended by the policy maker. With NH3-emmision

[7] "Inspire begint vandaag": Barry Woudenberg, Rob Peters en Marcel Hoogwout

maps that try to temper the intensive agriculture the zooming levels have to be very high to avoid battles in court. Here a few meters may cause severe economic damage. The Dutch regions do not want to lose this requirement of flexibility due to standardisation.

Sometime the visualisation of map layers for environmental purposes creates a legal challenge;

The feeding time of Goose in the Netherlands is six months at most. Then they fly back to the North. The map showing area's that are protected for the sake of those gooses should convey the temporal factor which is unusual in legal constraints mapping.

There are examples of cases where recreational agencies went to administrative courts for four times when they failed to recognise the intentions of the policymakers.

The biggest problem with European standardisation efforts like INSPIRE for Dutch regions are those that concern the harmonisation of calculating methods behind the layers. The Dutch regional authorities, perhaps the best Dutch experts in this area, are not entirely sure that all environmental indicators are calculated in the same way in Europe. This envisioned level of variety requires more explanation and enough administrative room for local interpretation. Even within the same country it takes time and effort to understand the different ways to calculate sound level measurements, visualisations and mappings. Water quality differs with temperature. Ecological complexity defies simple arithmetic's. There is a strong relationship between legislation and methods of calculation; Changes in calculation methods at the European level may cause changes in Dutch and regional legislation. INSPIRE may well be one of the directives with the highest impact on regional authorities[8].

This is both an argument for and against standardisation: if you insist on standardisation; then do it well. This line of argument plays a role for regional authorities in The Netherlands to become source owner of INSPIRE ANNEX map layers and to enter the European arena of standardisation.

4.1 Transparency and INSPIRE in a use case:

For the real life trial we have chosen to take the example of INPIRE annex III layer 11 and 18: area's with legal constraints, regulated area's and habitat/Natura2000 area's . The figure below depicts the Natura2000 map layer as it is now part of the information infrastructure of the region. The map is represented by yellow vertical stripes covering the entire Markermeer and IJssel Lake. In reality this is not the case. The Lake Markermeer requires numerous ecological measurements to save the area from further detoriation. The water management is suffering for slib that creates uninhabitable murky waters. Everyone agrees something has to be done. At the same time the already stressed ecological balance of the region is even more threatened by the obligation to host another 60.000 houses with accompanying recreational and transport infrastructure. Calling the region "locked under Natura2000" as indicated by the map layer is a simplification of our complex world that would bring people into the court room

Figure 6: Example of INSPIRE-type map layer

[8] As so defined by insiders within the Dutch community of regional authorities.

needlessly and too early. The groups involved in protection of birds and those involved in building the infrastructure have faced each other in court for the last 20 years. The chairman of the bird protection stated that the courtroom does not help his food chain to develop in a natural way. In reality the pressure groups require more insight in each other's negotiation space. Another reason to have dynamic maps fuelled by dynamic constraint analyses is the fact that no protected area is fixed in the way some maps tend to indicate. The small blue area's are protected breeding area's for a certain type of Duck for the ministerial website. In reality these regions may shift a few miles in each direction, depending on the best trade off in view of the NGO's involved. A more dynamic design of the relation between the map and the constraint it represent would actually support the legal cause much better.

Figure 7: Example of inspire type map layer

Inspire[9] has an explicit limitation to its scope as support mechanism for transparency: it states in article 13,1: " information services should not be published by means of inspire if this action jeopardizes the very environment it tries to protect. We argue here that statically defined map-layers powered by static constraints do exactly that. The do not support conflict resolution and negotiation but rather suggest more inflexibility and legal rigidity.

An important part of the interface is the geographical interface with Google earth and Google maps. The connection between the geographical data and the metadata that describes them can be seen in figure 2. The Map has a number of different regions which are of some type of land use. These regions have a particular geographical representations in the form of layers with polygons.

4.2 The " Vision" map: supporting on line conflict negotiations with legal maps

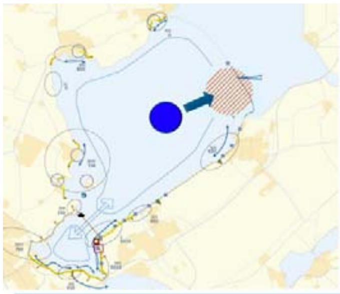


Figure 8: Conflict support-A

Early results with the test with the Flevoland regional development plan proves that questions like: return all contours on the map that fulfils the legal constraints "x" and "Y", but not "z" are indeed answerable. The challenge at the moment turns out to be technical and ergonomical. How does one provide such type of opportunity finding for the user?. Our research indicates that the most ergonomical lay-out for a use case would be, that the user drags the cursor (Blue, for op sailing area) and the surrounding area " tuns up" problem area's (striped red, in this case a 5000 HA swamp that is envisioned to be the ecological motor for the lake) and area's up for negotiation (orange). When the cursor is dragged further into the protected area it turns red to indicate the legal absolute negative. There are number of performance problems in Google Maps that need to be resolved with the interaction

Figure 8: Conflict support-B

[9] INSPIRE Work Programme Transposition Phase 2007 – 2009, d.d. 16 mei 2007.

between the cursor (blue) being dragged and constant category checks with the annotation engine.

4.3 Early results:

Figure 8: Conflict support-C

The interesting difference between a screen showing map layers already there and those that "turn up" while moving a qualified cursor (like the one seeking space for open sailing area's) is the affordance of opportunity finding in contrast with the annotation of an area that is " locked up"

The regional authorities call this functionality a "seeking area". They have created the legal term "seeking area" to obtain legal degrees of freedom in development plans that do not occur with fixed parameterisation.

The map layer turning up while approaching with the qualified cursor simulates the "seeking area" degree of freedom the authorities require to state intentions WITH a certain level of flexibility, like moving 2 km to the right or left if this helps with the original ecological goals. We call this the Simcity approach of Legal atlas.

5. CONCLUSIONS

The paper introduces maps as an effective support mechanism for conflict resolution sparked by legal data infrastructures like INSPIRE. We have covered the interaction between visualization of law using maps, it's problems with rigid interpretation and it's effect on real balancing of environmental issues with different societal interests. The approach described in this paper is similar to PPGIS systems to the extent that both use maps as visualization of issues that require balancing. In contrast, the discussion of normal PPGIS theory concentrates on the interface design or the forum capabilities of the tools, and so we suggest inclusion of the legal and semantic standardization possibilities in such platforms since it enables further query articulation and precision for citizens and businesses that seek negotiation space, opportunities and problems while the political decision process is underway.

The paper described a number of prototypes that have been iteratively built to make a case for this Legal Atlas approach together with the public authorities that govern complex issues that require conflict resolution rather than polarization. The participative and iterative building of the tools for solving problems supports a more appropriate design engineering oriented methodology for eGovernment research since statistical and analytical research would probably not have resulted in this Legal Map orientation for solutions. The paper described some Simcity type of ergonomics for easy access to negotiation space in potential conflicts. This approach will be tested in 2009.

INSPIRE is in interesting research area because in not only states what environmental indicator has to be measured; it for the first time also states how environmental indicators should be measured. We have highlighted some problems based on impact research among Dutch Provinces. The most important problem concern the harmonization of calculation methods of indicators with unwanted side effects. Further research should concentrate perhaps on European areas of legal and semantic implications for policy making and balancing environmental interests with economic and recreational interests in the context of SISE.

ACKNOWLEDGEMENTS

The authors wish to thank Mary Haselager and Kees Kersten of The province of Flevoland for their kind collaboration and providing us with all available data. We also thank Public-i (UK) for their design support.

REFERENCES

A. Boer and T. van Engers. A Knowledge Engineering Approach to Comparing Legislation. In *Proceedings of the Conference on Knowledge Management in Government*, Rhodes, Greece, Springer Verlag, 2003.

A. Boer, T. van Engers, and R. Winkels. Using Ontologies for Comparing and Harmonizing Legislation. In *Proceedings of the International Conference on Artificial Intelligence and Law (ICAIL)*, Edinburgh (UK), ACM Press. 2003.

A. Boer, R. Hoekstra, R. Winkels, T. van Engers, and F. Willaert. META*lex*: Legislation in XML. In T. Bench-Capon, Aspassia Daskalopulu, and R.G.F. Winkels, editors, *Legal Knowledge and Information Systems (Jurix 2002)*, pages 1-10, Amsterdam, IOS Press, 2002.

A. Boer,R. Peters and T. van Engers. Separating Law from Geography in GIS-based eGovernment Services; Journal for AI and Law, 2007.

Craig, et al. *Community Participation and Geographic Information Systems*. London: Taylor and Francis, 2002.

E. Glasséé, T.M. van Engers, A. Jacobs, POWER: An Integrated Method for Legislation and Regulations from their Design to their Use in E-government Services and Law Enforcement., In M-F. Moens (ed.), *Digitale Wetgeving, Digital Legislation*, 175- 204, Brugge, Die Keure, 2003.

E de Maat, T.M. van Engers, Mission Impossible?: Automated Norm Analysis of Legal Texts, In D. Bourcier (ed), *Legal Knowledge and Information systems, Jurix2003: the sixteenth Annual conferenece,* ISBN 1 58603 398 0, Amsterdam, IOS Press, 2003.

R. Peters and T. van Engers. The legal atlas: Map-based navigation and accessibility of legal knowledge sources. Knowledge Management in Electronic Government; 5th IFIP International Working Conference, KMGov 2004, Krems Austria, Springer Verlag, 212– 220, 2004.

F. Wilson and R. Peters. Mapping the law: Knowledge support for business development enquiry. In Proceedings of the eChallenges 2004 Conference, IST Programme, Vienna, 2004.

EC-INSPIRE: Spatial Data Infrastructures in Europe: State of play Spring 2003 commissioned by EUROSTAT & DGENV, K.U. Leuven, 2003.

Cartwright, W., Crampton, J., Gartner, G., Miller, S., Mitchell, K., Siekierska, E., et al. Geospatial Information Visualization User Interface Issues. *Cartography and Geographic Information Science , Vol. 28 (No. 1)* 2000.

Malczewski, J., GIS and Multicriteria Decision Analysis. Wiley and Sons, J. 1999.

van de Ven, S., Hoekstra, R., & Winkels, R., MetaVex: Regulation Drafting meets the Semantic Web. *SW4Law Workshop* , 27-32, 2007.

Winkels, R., & Hupkes, E., Legal Atlas: Access to Legal Sources through Maps. *Proceedings on the 11th International Conference on Artificial Intelligence and Law (ICAIL 2007)* . Amsterdam: IAAIL, ACM, 2007.

Human Factors Research Group. SUMI. *Software Usability Measurement Inventory*. University College, Cork, 1994.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Access to Environmental information in a Shared Environmental Information System

Thomas Pick

Niedersächsisches Ministerium für Umwelt und Klimaschutz
Hannover, Germany (thomas.pick@mu.niedersachsen.de)

Abstract: The rapidly changing natural environment asks for appropriate political, legal and technical measures in order to continue to facilitate the resources necessary for a life in comfort and safety in the European Community. The availability, accessibility, accuracy and accountability of environmental data and information is the prime requisite for mastering the challenges that lie ahead. The European Community has addressed this notion with a series of legislation that aims at making environmental information available to the public at minimum costs.

The Directive on public access to environmental information (2003/4/EC; AEI-Directive), together with a suit of further cross-domain legislation (e.g. the Directive on Reuse of Public Sector Information; Directive on Public Participation and the INSPIRE Directive) forms a cornerstone of this policy. It requires member states to progressively make available environmental information and disseminate it to the public in order to achieve the widest possible systematic availability and dissemination. It grants a general right of access to information to any person, not limited to EU citizenship and thereby constitutes one legal building block to what has recently become known as the shared environmental information system, SEIS.

The AEI-Directive has been in effect for 6 years. In the context of the EU FP7 Project "ICT-ENSURE" we have conducted a survey on the state of implementation of the AEI-Directive in member states. This paper summarizes the preliminary results from this study.

The study is based on an internet survey. The member state specific results from this study have been summarized in an individual report sheet for each country. The thus prepared report cards were then been sent to the publication departments of the Ministries of Environment of the EU 27 for review, correction and completion. Up to today 10 countries have responded and sent back the review report cards.

The study shows that all member states use the internet as the main distribution tool for environmental information. Legislation for free access to information is in effect in the majority of the countries. More than half have implemented the AEI-Directive in discreet legislation. Except for one country, member states have decided not to develop special information systems for the technical implementation of the directive.

Keywords: SEIS, SISE, Environmental Information Directive

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Environmental Geoinformatics for Emergency Management

J. Ráček[a], T. Ludík[a], P. Kubíček[b], E. Mulíčková[b], G. Šafr[b]

[a] *Faculty of Informatics, Masaryk University, Botanická 68a, 602 00 Brno, Czech Republic (racek@fi.muni.cz, xludik2@fi.muni.cz)*

[b] *Department of Geography, Faculty of Science, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic (kubicek@geogr.muni.cz, evamulickova@mail.muni.cz, gsafr@mail.muni.cz)*

Abstract: Presented paper illustrates issues of geodata content definition and visualization within emergency management and combines them with the process modelling. Because of the overall complexity the proposed solution is described with a limited parameters using specific scenario dealing with the traffic accident and transportation of dangerous chemical substances. First of all the scenario domain description is given. It includes the rescue units intervention at an accident with dangerous substances, commander intervention related activities and their cooperation on tactical and operational levels. Activities mentioned above are further analysed with using document "Typical activities of the IRS components at the common intervention". Process is viewed, structured and presented by use case diagrams. It identifies processes on the highest level of process hierarchy. Then the process analysis and modelling of the emergency activities are decomposed in top-down manner and the XPDL (XML Process Definition Language) is used as the primary process definition format. Based on the analysis of the action (e.g. intervention organizing) and types of situation (e.g. accident of vehicle with dangerous substances) in emergency management it is possible to define context-aware scenario in which geographic information supports decision making.

Keywords: Emergency Management; Environmental Geodata; Use Case; Process Model.

1 INTRODUCTION

The field of emergency management usually divides any situation into four components (phases) that roughly corresponds to the before – called *preparedness*, during – called *response*, and after phase (*recovery*) of any particular event. The last but not least *mitigation* phase is concerned with minimizing the effects of possible disasters by mandating policies and regulations that will lessen the effects of hypothetical disaster.

Many approaches to develop a geoinformation driven system for disasters have concentrated on the response phase. This phase is dealing with the immediate aftermath of an emergency including the mobilization of relief agencies, the delivery of aid and the provision of emergency medical care. 0 speculates that this is due to the often publicized failures of emergency services in this phase. Also in the Czech Republic there exists the common will for the development of an effective information support of intervention commander. This is especially valid for extensive and complicated emergency actions where general information and geoinformation support is a must. Direct visualization and update of geoinformation in the filed should simplify the decision making process of intervention commander and rises the quality of work of Integrated Rescue System at all levels.

Presented work is based on the scenario "Transportation of dangerous chemical substances" described by 0 which is focused on verification of the dynamic geovisualisation procedures in the selected crisis situation. The scenario was proposed in 2007 to verify some theoretical geovisualisation approaches, and at the same time, to verify the functionality of the communication and information systems that are designed as the emergency system components. The presented contribution further develops the proposed system by integrating of 3 interdisciplinary fields: *emergency management*, *process modelling* and *adaptive cartographic visualization*.

2 EMERGENCY MANAGEMENT IN THE CZECH REPUBLIC

Civil Protection is a complex of the prepared measures that are practically realized at Extraordinary Events (EE) and Crisis Situations (CS). These measures are executed by components of the Integrated Rescue System (IRS).

IRS is legally specialized, open system of coordination, cooperation and modelled cooperation procedures (Typical activities). It is predestined for settlement of the day to day events, natural and anthropogenic disasters.

Fire Rescue Corps of the Czech Republic (FRC) and Fire Prevention Units (FPU), Police of the Czech Republic (PCR) and Medical Rescue Service (MRS) are the basic components of the IRS. Other IRS components include the local emergency services (according to their profession), municipal/town police, health care institutions, Army of the Czech Republic, civil associations, eventually legal and natural persons.

Coordination of the Rescue and Liquidation Works (RaLW) is a procedure realized by the IRS corresponding organs. By applying managing functions and IRS components and by using accessible resources of Forces and Means (FaM) it allows mastering of existing extraordinary situation. Coordination is done on three levels – *tactical*, *operational* and *strategical* 0.

2.1 Tactical level

It includes coordination of the RaLW at the place of intervention of IRS components and cooperation management of IRS components.. According to the Extraordinary Event consequence, intervention commander proclaim corresponding Level of Alert, which predetermine needs of the Forces and Means (FaM) for Rescue and Liquidation Works (RaLW). At simply cases, the intervention commander is able to coordinate the FaM alone, respectively with cad of intervention commander. At complicated cases, which require time consuming and complex cooperation, the intervention commander nominates the staff of intervention commander (staff). As staff members there are nominated the IRS components leaders, eventually experts or assistants with which units cooperate on place of intervention. It is also possible, in the case of too complex or large-scale intervention, to organize individual place of intervention sectors and nominate the sector commanders.

In any case the intervention commander organizes the RaLW on bases of coordination principle. Based on the consultation with IRS component leaders he/she determines the RaLW of FaM common procedure that should follow the "Typical activities of the IRS components at the common intervention".

2.2 Operational level

At this level permanent coordination and cooperation is taking place - it includes operational centres of the IRS basic components (FRC, PCR, FRS) and dispatching centres, standing services, oversight centres of some other IRS components (especially of distributive and emergency services). Cooperation is done both within and between individual IRS components.

Operational cooperation consists in consistent and common activity of IRS components and grouping of Forces and Means at the place (places) of intervention at the area affected by Extraordinary Events. Operational cooperation within the Rescue and Liquidation Works at common activity of the IRS is realized and managed from the operational control level by IRS Operational and Informational Centre (OIC) which is generally identical to Fire Rescue Corps of the Czech Republic OIC (FRC OIC) of the region. Main activities of the IRS OIC covers: fulfilling the intervention commander requirements, fulfilling tasks given by authorities empowered to coordinate RaLW (General Management of the Czech Republic FRC, FRS of the region, region commissioner, mayor of the municipality with extended powers), notification of the affected IRS components and administration bodies, FaM call up and appointment, realizing citizens warning in the endangered area, etc. Main task of the IRS OIC is also to secure uninterrupted support of the crisis staffs activities and exchange of information.

To manage coordination and cooperation of the IRS components on the operational level, IRS documentation is used. Among others it includes IRS alert plan, Emergency plan of the region, External alert plans, Water sources survey and also document called “Typical activities of the IRS components at the common intervention” which models activity of the IRS component at Rescue and Liquidation Works with regard to the character of Extraordinary Events.

2.3 Strategical level

Strategical level of the RaLW management, under required conditions for coordination taking-over (EE extent, highest alarm stage, coordination demand) are realized by standing or temporary coordinating authorities of the administration, region commissioners and Ministry of Interior – General Management of the Czech Republic Fire Rescue Corps. Temporary coordinating authorities for RaLW are the Crisis Staff of the Municipality with Extended Powers and Regions, Ministry of Interior – General Management of the Czech Republic Fire Rescue Corps Staff. Standing coordinating authorities for RaLW are the IRS OICs and IRS OIC of the General management of the Czech Republic Fire Rescue Corps.

3 PROCESS MODELLING

A process modelling is used for depiction of a sequence of operations, declared as work of a person, work of a simple or complex mechanism, work of a group of persons, work of an organization of staff, or machines. Process modelling may be seen as any abstraction of real work. For control purposes, process model (workflow) may be a view on real work under a chosen aspect, thus serving as a virtual representation of actual work. The flow being described often refers to a document that is being transferred from one step to another.

A process model is a model to represent real work for further assessment, e.g., for describing a reliably repeatable sequence of operations. More abstractly, a process model is a pattern of activity enabled by a systematic organization of resources, defined roles and mass, energy and information flows, into a work process that can be documented and learned. Process models are designed to achieve processing intents of some sort, such as physical transformation, service provision, or information processing.

Process modelling concepts are closely related to other concepts used to describe organizational structure, such as silos, functions, teams, projects, policies and hierarchies. Workflows may be viewed as one primitive building block of organizations.

To solve emergency situations in the most effective ways it is necessary to abandon function management, which is focused on the knowledge of individual persons. In this system the tasks are assigned to individual people who execute them. According to the character of these tasks people are divided into separate function units. Process management is a better option. This way of management is focused on work results. Work is not done separately in individual function units but it flows through them. When this kind of management is used, improvements are reached by optimization of the workflow and by simplifying it. Process management leads to more effective coordination of work and also to decrease of fault occurrence.

4 CASE STUDY

Research presented in this paper focuses on the event "Accident of vehicle transporting dangerous substance". Chapter 4.1 describes this event in complex manner as use case diagram. In next sessions case study is defined - organizing the intervention within event "Accident of vehicle transporting dangerous substance". Chapter 4.2 describes this case study as process model and chapter 4.3 CRUD matrix that defines operations on map features in the process is shown.

4.1 Use case diagram

Before focusing on the process maps it is necessary to understand the event "Accident of vehicle transporting dangerous substance" as a complex 0. It is possible to create a good complex view with the help of the use cases when the event was applied. It is represented by the UML (Unified Modelling Language) use case diagram. The main purpose of the use case diagram is to find and document the modelled system requirements 0. The first stage of creation of this diagram is to specify the border of the modelled system. In this case, the border is defined quite clearly by the Czech Fire and Rescue Act. Everything else is considered to be surroundings of the system.

Next, an actor list should be created. It is a list containing different roles which are assigned to persons or subjects that use the modelled system. In order to create the complete actor list, it was necessary to analyse all the activities within "Accident of vehicle transporting dangerous substance". When analysing these activities the most important questions were: "Who or what uses the system?" and "Who or what communicates with the system?"

The identified actors based on "Typical activities of the IRS components at the common intervention" are (Fig. 1):

- Intervention commander,
- Integrated rescue system,
- Police,
- Fire rescue corps,
- Medical rescue service,
- Operational and informational centre,
- Fire Prevention Units,
- Public administration bodies,
- Public health protection authority.

Having understood the roles of the individual actors, it is possible to start creation of use cases. A use case is perceived as specification of the sequence of activities, including variable sequences or fault sequences, that the system or subsystem can execute through interaction with external actors. There are five use cases in main use case diagram. These use cases were designed on basis of "*Typical activities of the IRS components at the common intervention*" - similar actions were grouped to form a use case (Fig. 1).

- Evaluation of intervention,
- Organization of intervention,
- Decontamination of intervention,
- Coordination of health care,
- Close of intervention.

After modelling the use case diagram, it is possible to start creation of the specifications for the individual use cases. These specifications are described as processes 0 (using process maps) incorporating and defining the activity sequences in the particular directives.

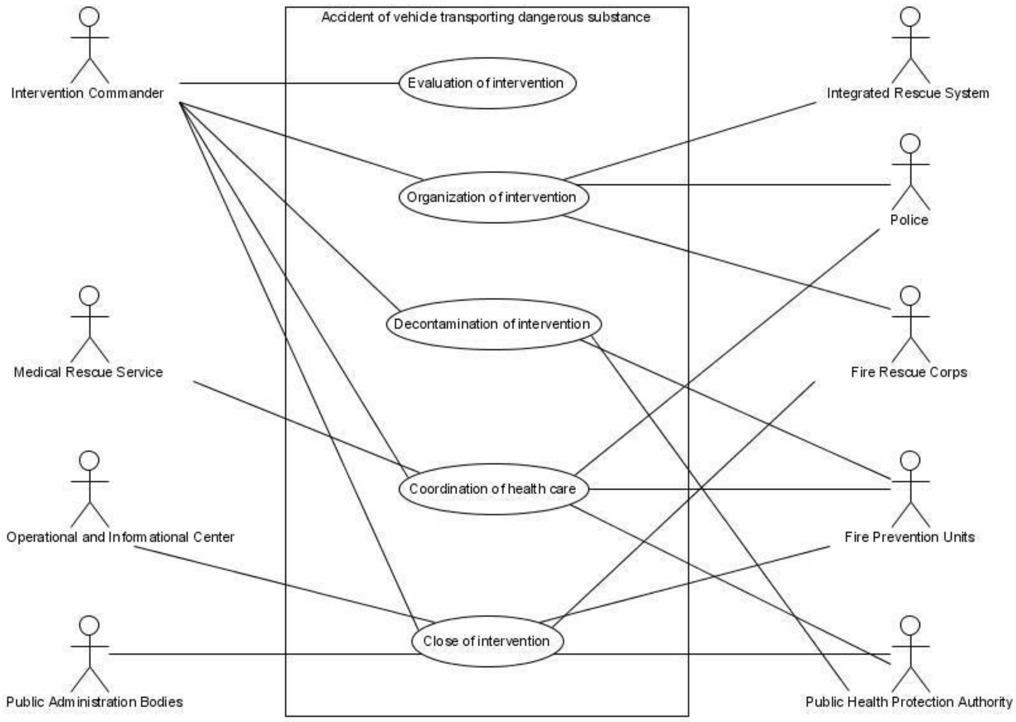


Figure 1: The main use case diagram of accident of vehicle transporting dangerous substance

4.2 Process maps

A process is a set of activities arranged in parts which creates in a repeatable way a required output on the base of one or more inputs. Specifications of the individual processes consist of the following entities:

- Process (Workflow Process Definition),
- Activity (Workflow Process Activity),
- Transition (Transition Information),
- Participant (Workflow Participant Specification),
- Application (Workflow Application Declaration),
- Data (Workflow Relevant Data).

Mutual relations among the entities are demonstrated in the process meta-model 0. To illustrate a process map the use case called "*Organization of intervention*" is processed. This directive controlled by Intervention commander consists of ten activities:

1. Determination of intervention organization,
2. Determination of hot and war endangered zones,
3. Determination of exposure zone,
4. Determination of entry area
5. Determination of rare area
6. No entrance for unauthorized person,
7. Call to fire rescue corps,
8. Call to police,
9. Determination of enclosure of exposure zone and its regime
10. Determination sectors, section and its commanders.

When we want to represent this situation with a particular process map, we have to use two routing actions, with AND-Split and AND-Join and also three activity blocks for the main processes. Of course, it is necessary to add marks for the start and the end of the process. In this way the process map of organization of intervention is created (Fig. 2).

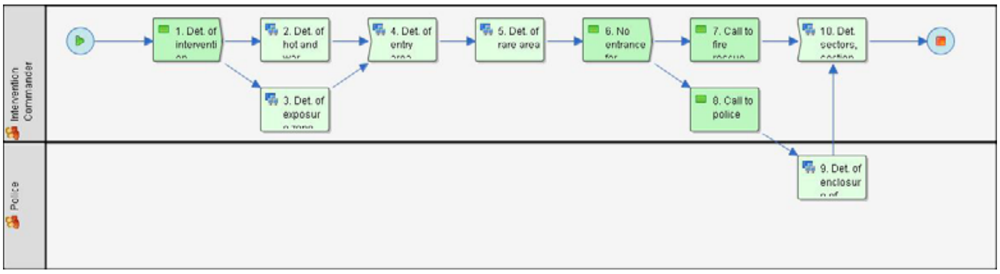


Figure 2: The process map of organisation intervention.

All modelled processes (process maps) are transformed XPDL (XML Process Definition Language) format (Fig. 3). It is Workflow Management Coalition standard, which is trying to create a uniform format for storage of the modelled processes. Nowadays this standard is supported by 70 software products. The main goal of this standard is to create a process definition in such a format, that the models can be transferred to different modelling tools or automatically processed by the workflow management system 0.

```
<Activity Id="newpkg1_wp1_act4" Name="4. Det. of entry area">
    <Implementation>
        <Tool Id="Visualization Server" Type="APPLICATION">
            <ActualParameters>
                <ActualParameter>BASETOPO</ActualParameter>
                <ActualParameter>OBJECT-HAZARD</ActualParameter>
                <ActualParameter>OBJECT-IN_DANGER</ActualParameter>
                <ActualParameter>LOCAL</ActualParameter>
                <ActualParameter>ZONE-A</ActualParameter>
                <ActualParameter>ZONE-B</ActualParameter>
                <ActualParameter>ZONE-C</ActualParameter>
                <ActualParameter>B-ENTRY</ActualParameter>
            </ActualParameters>
        </Tool>
    </Implementation>
    <Performer>Intervention Commander</Performer>
</Activity>
```

Figure 3: Emergency gate determination activity in XPDL. Parameters in capital letters represent name of map feature which is described further in the text.

4.3 CRUD matrix

The process formalisation in XPDL form is enabled by geoinformation analysis of the case study. It specifies which geoinformation is needed to be used and finds an appropriate way of their representation and processing. Subsequently, the processes and geodata is rebalanced using CRUD matrix. The matrix shows which activities are executed in the process and which geoinformation is needed. The lines of the matrix represent geodata items associated with map symbols; the columns represent particular process activities. In the particular matrix fields there are the following operations: create (C), read (R), update (U) and delete (D), which are executed by the process with the respective geoinformation. CRUD matrix is illustrated in next session in figure 4.

It is possible to distinguish 2 kinds of operation within the use case that user can apply on geodata. These operations correspond to those given in CRUD matrix. *READ* - simple view of information from server - this operation can be applied on data available over the Internet, no matter if stored in DB (e.g. data from different providers), data provided as web services, data resulting from analysis on server or data once created in-situ and saved on server. *CREATE* - information entry - editing of data on client side. Other operations

(*UPDATE*, *DELETE*) are not taken into account since they are covered by another use cases.

	1. Determination of intervention organization	2. Determination of hot and war endangered zones	3. Determination of exposure zone	4. Determination of entry area	5. Determination of rare area	6. No entrance for unauthorized person	7. Call to fire rescue corps	8. Call to police	9. Determination of enclosure of exposure zone and its regime	10. Determination sectors, section and its commanders
OBJECT-HAZARD		R	R	R	R				R	R
OBJECT-IN_DANGER		R	R	R	R				R	R
LOCAL		R	R	R	R				R	R
ZONE-A		R	R	R	R				R	R
ZONE-B		R	R	R	R				R	R
ZONE-C			R	R	R				R	R
A-TECHNICAL										C
B-DECONTAM										C
B-ENTRY				C	R				R	R
B-COMMANDER										C
B-HEALTH										C
C-RANK					C				R	R
C-ACCOUTRE					C				R	R
ENCLOSE									C	R
AIR_RESCUE										C
EVAC_STAND										C
BASETOPO		R	R	R	R				R	R

Figure 4: CRUD matrix of use case „organizing of intervention". In rows - map feature (further explained in text) and type of allowed operation (R-read, C-create), in columns - activities within use case

5 CONSLUSION

Geographic visualization within activities of Integrated Rescue System is becoming a must. It can significantly support decision making on all levels of emergency system management - e.g. optimisation of Forces and Means by Rescue and Liquidation Works, solutions in crisis situations with exposure of human lives, possessions, unites of Critical Infrastructure etc.

The presented article deals with the methods of geoinformation support automation within the emergency management. The authors are aware of other attitudes proposed and often used in this area like semantics driven or ontology development (e.g. 0,0). However, authors tried to build the support in a different manner starting with a simple use case, identifying main obstacles and possible bottle necks for a more extensive implementation.

This interdisciplinary research focuses on connecting process modelling and adaptive cartographic visualization in emergency management. Such complex solution has two main advantages - it allows user to get the right information in the right moment. The right information is determined mainly by user needs for specific action in specific situation. Based on the professional and cultural background of the user and device used such information can be transmit to user in user friendly manner (i.e. is adaptive visualized). The right moment is given by analysis of task flows within the action that user is responsible for (i.e. process map). By process map it is determined which task needs support of spatial information and what operations are allowed on it.

Proposed solution enables machine understandable support and process automation within emergency management.

ACKNOWLEDGEMENTS

The authors wish to thank the Ministry of Environment of the Czech Republic for the financial support of the project no. SP/4h4/168/07 and the Ministry of Education, Youth and Sports of the Czech Republic for the financial support of the research plan no. MSM0021622418.

REFERENCES

Fiala, J. and Ministr, J., The Model of Process Framework in Czech Public Government, Change Management, Proceedings of the 26th International Conference on Organizational Science Development. University of Maribor, Portorož, 2007.

Halliday, Q., Disaster Management, Ontology and the Semantic Web, UCLA Department of Information Studies, 2007.

Hollingswort, D., Terminology & Glossary, Workflow Management Coalition, 1999.

Klien, E., Lutz, M. and Kuhn, W., Ontology-based Discovery of Geographic Information Services – An Application in Disaster Management, Computers, Environment and Urban Systems (CEUS), 2005.

Ludík, T. and Ráček, J., Procesná analýza krízového riadenia a jej využitie pri tvorbe aplikačných ontológií, Riešenie krízových situácií v špecifickom prostredí, Žilinská univerzita, Žilina , 2008.

Šafr, G., Struktura a legislativa Integrovaného záchranného systému (IZS), koordinace a návaznost činností složek IZS, mimořádné události a krizové situace, Jihočeská univerzita, České Budějovice, 2008.

Staněk, K., Konečný, M. and Friedmannová, L., An Adaptive Cartographic Visualization for Support of the Crisis Management, XXIII International Cartographic Conference - Cartography for everyone and for you, Roskartografija, Moscow, 2007.

Talhofer, V., Brázdilová, J., Kubíček, P. and Svatoňová, H., Transport of Dangerous Chemical Substances and its Modelling In Sborník 10th Conference on Geographic Information Science, AGILE 2007, Aalborg, 2007.

Tanasescu, V., Gugliotta, A., Domingue, J., Villarías, L., Davies, R., Rowlatt, M., Richardson, M. and Stincic, S., Geospatial Data Integration with Semantic Web Services: the eMerges Approach, The Geospatial Web, eds. Arno Scharl, Klaus Tochtermann, Springer. 2007.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pilmmann, I.Holoubek, T. Bandholz (eds.)
Masaryk University, 2009

Review of Croatian ICT Research Activities on Sharing Environmental Information

Ivica Ruzic and **Jadranka Pecar-Ilic**

Division for Marine and Environmental Research, Rudjer Boskovic Institute, Zagreb HR 10 000, Croatia (ruzic@irb.hr, pecar@irb.hr)

Abstract: The paper gives a review of research and development activities of Croatia, where application of information and communication technologies significantly improves environmental information management and reporting. The most important ICT based research activities on several environmental projects either at the national or international level are presented and discussed. For example, some interesting results of the Adriatic Sea investigations are mentioned. Within the Environmental Programme for the Danube River Basin, Croatia contributes to efficient data management by developing Web GIS for the Transnational Monitoring Network. Major activities on implementation of river information services at the pan-European and national levels are also described. Specialized information system for surveillance of the Croatian segment of the Danube River enables temporal and spatial presentation of waterway data. Information system architecture, based on the Web-GIS-DBMS integration, is presented. The very important task of sharing environmental information is the responsibility of the Croatian Environment Agency. An integrated environmental information system has been recently developed by using modern information technologies and communication standards in line with the European requirements.

Keywords: Adriatic Sea; Danube River Basin; Environmental information system; GIS; River information services.

1. INTRODUCTION

During the difficult war period from 1991, when Croatia became independent, until 1995, very little effort was invested into ICT based developments in all areas and so also in environmental protection. Later, development of more sophisticated environmental information systems was initiated, first within different international research and development programmes in which Croatia was able to participate. The most important results were achieved within the Environmental Programme for the Danube River Basin (EPDRB). Since 2006, the Croatian Government has made great efforts to enhance such activities at the national level by fostering long-term investigations of environmental protection and ICT as well as short-term investigations of transport safety and its influence on the environment. Of special interest is the development of Geographic Information Systems (GIS), web applications and services, in-situ measurement systems and monitoring of the environment based on networking of sensors and communication systems. In addition, in 2008 the Croatian Government passed a regulation committing the Croatian Environment Agency (CEA) to set up an integrated environmental information system aimed at integrating environmental data and reports at both national and European levels.

Since 1995, our research and development projects have included ICT application to environmental protection and river navigation surveillance. The main focus was the development of specialized information systems, web applications for temporal and spatial presentations and services necessary to provide efficient management of complex data and

processes in investigations of water quality, natural and anthropogenic pollution sources, risk assessment and particularly the influence of river navigation on the waterways.
This paper gives a brief review of relevant research and development activities of Croatia from the very beginning to the present day. As we have been personally engaged on international environmental programs as well as scientific research projects at the national level, we will discuses the most significant activities and results.

2. INVESTIGATIONS OF THE ADRIATIC SEA

The first important international investigations concerning the Adriatic Sea were carried out by the Coordinated Mediterranean Research and Monitoring Programme (MED POL). Long-term objectives of the MED POL-Phase II were to prevent, abate and combat pollution of the Mediterranean Sea area and to protect and enhance the marine environment of the area. Specific objectives were, among others, to provide information that could be used in formulating environmentally sound national, bilateral and multilateral management decisions essential for the continuous socioeconomic development of the Mediterranean region on a sustainable basis, as specified by UNEP [1994]. Later on, UNEP/MAP created several specific Regional Activity Centres (RACs), such as the Priority Action Programme (PAP) based in Split working on integrated coastal planning and management. Of special importance is also the relatively new RAC/INFO based in Palermo, organized to build MAP's information and communication capacity. Cooperation is planned between UNEP/MAP (InfoMAP proposed a shared information system in the context of the Barcelona Convention) and the European Environmental Agency (EEA). EEA with its partnership network EIONET is developing a shared environmental information system as specified by EEA, UNEP/MAP [2007].
At the national level, the very first scientific effort was made in 1995 to design the conceptual model of an integral referral information system for the Adriatic Sea investigations in Croatia. This comprehensive description of organizational and functional aspects of the performed investigations includes meta information about projects, responsible institutions, stations, samples, measuring methods and parameters, as shown by Pecar-Ilic et al. [1997]. Since 1998 onwards, the Croatian National Monitoring Programme–Jadran enables systematic research on the Adriatic Sea as a basis for sustainable development of the Republic of Croatia, as described by Project Jadran [2009]. The project is sponsored by the Croatian Government and research and development activities are carried out by several scientific institutions. Long-term goals involve systematic ecosystem research of the whole national marine territory, especially estuaries, strategic marine research, development and application of new technologies and instruments and improvement and extension of cooperation between different states and research projects. ICT based developments include remote sensing methods and automatic measurement systems with in situ data sending, as shown by Dadic et al. [2007].

2. INVESTIGATIONS OF THE DANUBE RIVER BASIN

2. 1 Results of the Environmental Programme for the Danube River Basin

A very important step in the development of environmental information systems in Croatia was made by setting up the Environmental Programme for the Danube River Basin (EPDRB). It was initiated in 1991 and lasted until the International Commission for Protection of the Danube River (ICPDR) was established in 1998. The main goal of this Programme was to strengthen the sustainability of water quality management in the Danube Basin. For this purpose, several permanent expert groups were established, as described by Lack [1997], such as the Monitoring Laboratory and Information Management Expert Group (MLIM), Emissions Expert Group (EMIS) and Accidental Emergency Warning System Expert Group (AEWS). Croatia took an active part in all these activities.

Our research was initiated in 1996 within the MLIM expert group and we were responsible for information management. Investigations were conducted in cooperation with the company "Croatian Waters". We produced the first web pages for the Trans-National Monitoring Network, TNMN, as presented by Ruzic et al. [1997]. From 1997 to 2000, we played the role of database administrators for the entire TNMN. We also gave the first experimental presentation of web pages using wireless communication, as described by Ruzic et al. [1997], and in 1999 the first experiments of integrating Geographic Information Systems (GIS) and databases with automatic data presentation on the web, as shown by Ruzic, Pecar-Ilic [2000].

2. 2 Development of Web GIS and applications

From 1998 to 1999, we contributed to the development of the Danube Information System for ICPDR (DANUBIS) collaborating with the International Society for Environmental Protection (ISEP) from Vienna, as shown by UNDP/GEF, ISEP [1999]. The system offers internal information to ICPDR and its working bodies and information distribution to the public. For this purpose, a special web site has been established with links to national web pages, which make the best use of available information sources and communication tools to achieve the objective of sustainable development in the Danube region, as described by UNDP/GEF, ISEP [1999]. In order to enable timely environmental information, the architecture of the specialized information system should include functionalities of both the Relational Database Management System (RDBMS) and Geographic Information Systems as elaborated by Pecar-Ilic [2001].

Our research experience has shown that spatial distribution of monitoring stations plays a significant role in the process of relevant data management. For example, a database could be used to manage all the necessary information dealing with water quality and quantity monitoring. GIS could be used for efficient processing of spatial data. Information system based on the performed WEB-GIS-DBMS integration could enable efficient temporal and spatial presentations of environmental data, as presented by Pecar-Ilic, Ruzic [2001]. The digital map of the Danube River Basin containing monitoring stations included in Trans National Monitoring Network (TNMN), together with an adequate legend presenting active layers, is shown in Figure 1.

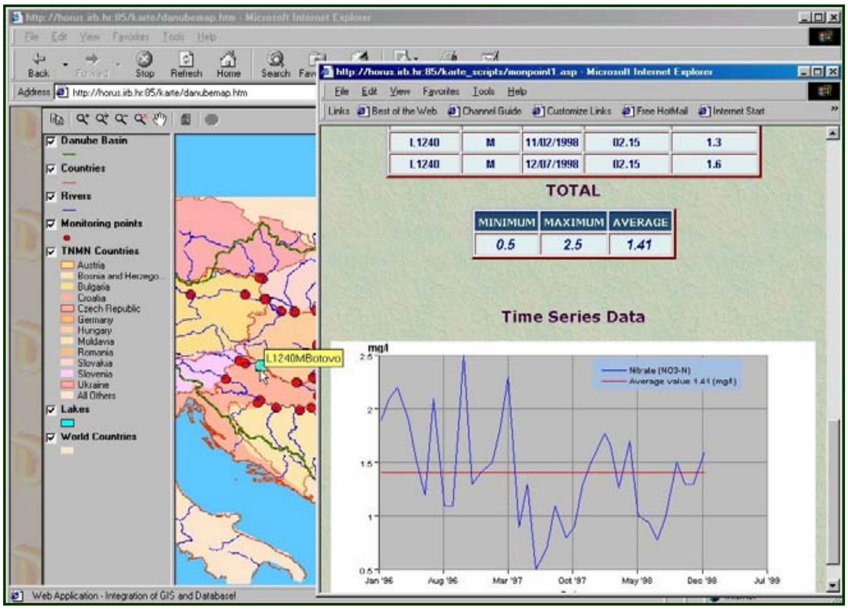


Figure 1. Example of the generated time series data for *Nitrate (NO_3-N)* at the *Botovo (L1240M)* monitoring station in the Trans National Monitoring Network.

Finally, the Danube River Basin Geographic Information System was developed within the UNDP/GEF Danube Regional Project as a core tool for efficient river basin management according to the Water Framework Directive. GIS provides a centralized platform for exchanging, harmonizing and viewing geo-information and related issues. Further development towards decentralized system architecture, and its connection to and interoperability with other tools and systems are planned for the future, as shown by UNDP/GEF [2006].

3. IMPLEMENTATION OF RIVER INFORMATION SERVICES

3.1 Croatian activities at the pan-European and national levels

Based on directive *2005/44/EC* of the European Parliament and the Council (so called *RIS Directive*), RIS guidelines describe the principles and general requirements for planning, implementing and operational use of river information services and related systems, as specified by European Commission [2005]. The “GIS Forum Danube” is regarded as a perfect platform for managing RIS oriented projects such as Consortium Operational Management Platform River Information Services (COMPRIS) and Data Warehouse for the Danube Waterway (D4D), as shown in GIS Forum Danube [2004]. Croatia also actively participates in the activities of the “GIS Forum Danube”, as well as in expert groups of these two projects as described by Pecar-Ilic, Ruzic [2006]. Their main goals are to define the concept of common Danube waterway related data. For example, an important objective is to implement the International Association of Lighthouse Authorities (IALA) beacons for the establishment of differential Global Positioning System (dGPS), as shown by Pfliegl [2004]. Inland Electronic Chart Display and Information System (ECIDS) standard is used for accurate digital map production.

To involve Croatia in the RIS system development, the CROatian River Information Services (CRORIS) project was started in 2003, as described by CRUP [2009], using expert and institutional cooperation with COMPRIS.

3.2 Architecture of the River Information System (River IS)

Since 2000, we have been conducting research activities for the Croatian national project which will improve waterway related data management in Croatia. For that purpose, a specialized geographic information system for surveillance of the Croatian segment of the Danube River (i.e., *River IS*) started to be developed and in 2003 it became a part of the CRORIS project. *River IS* possesses the main characteristics of both the Geographic Information System (GIS) and the Data Base Management System (DBMS), as elaborated by Pecar-Ilic [2001] and Pecar-Ilic et al. [2002]. The emphasis is placed on sharing relevant and timely information among different River IS user groups as well as on efficient data management provided by this information system. The main components (packages or subsystems) of the River IS architecture were formally described at the highest level by the standard Unified Modeling Language (UML).

A high level view of the Croatian *River IS* shows provision of information to various user groups of the River Information System. Developed Web applications, as shown in Figure 2, enable several functionalities, as described by Pecar-Ilic, Ruzic [2004]. Employing XML as the format of exchanged data, as well as some of the supporting XML technologies in developed web applications, helped establish a standard for exchanging data between CRORIS, its partners, and the regulatory authorities, as presented by Bec et al. [2006].

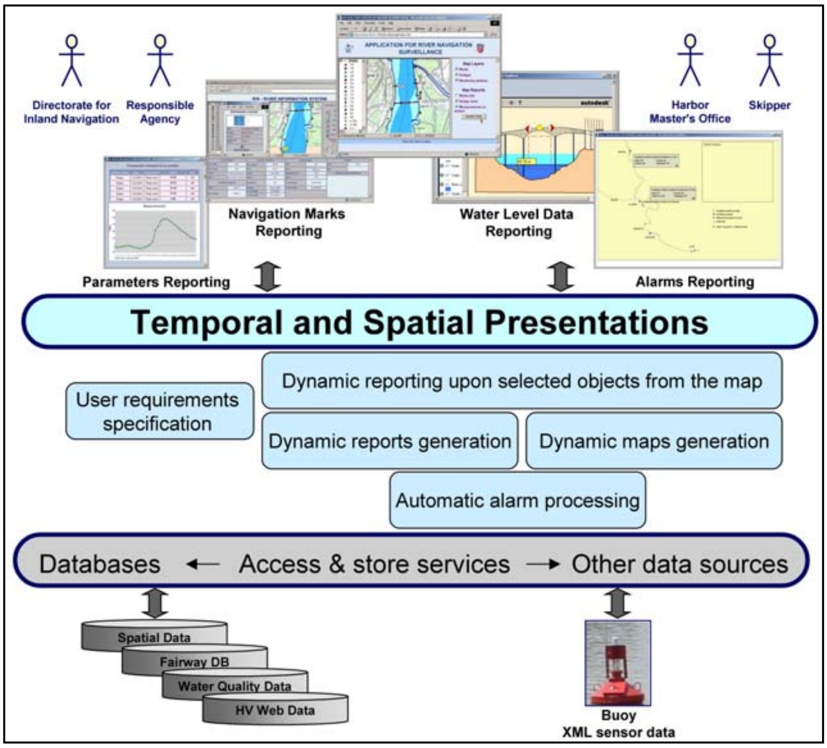


Figure 2. A high level view of the *River IS* architecture.

4. DEVELOPMENT OF THE ENVIRONMENTAL INFORMATION SYSTEM

4.1 Objectives of the Croatian Environment Agency

In 2000, the need was recognized for a comprehensive environmental data system at the national level comprising data on the current state of the environment, performed and planned monitoring, pollution and contamination, pollution and load trends, environmental projects implemented or planned in Croatia. The Croatian Government established the Croatian Environment Agency (CEA) in 2002 as an independent public institution. Its main objectives are collecting data from different sources, data integration and processing and environmental reporting at the national level, their processing, maintaining the environmental database, environmental monitoring and reporting. CEA has established close collaboration with the European Environment Agency (EEA) and has joined the European Environment Information and Observation Network (EIONET).

As described by CEA [2009], important projects include establishment of National Reference Centres (NRCs), setting up the national Environmental Information System (EIS), preparation of the State of the Environment Report, creation of the Croatian Environmental Projects Database, Risk Installations Database and Waste Management System. In addition, high priority is also given to the development of the Croatian Soil Monitoring Programme and the special Concept of Reporting on Water (CROW).

4.2 Environmental Information System

The objective of setting up an integrated EIS was reached through intensive collaboration with a number of institutions and organizations engaged on data collecting and processing for years. They are in possession of certain databases related to different components of the environment.

Therefore, EIS is seen as an integrated system of several information themes concerning different components of the environment, as described by CEA [2009]. The main themes include the sea, soil, air, water, waste and biodiversity. The Environmental Information System is developed by using modern information technologies and communication standards in line with the European requirements, and by developing the environmental data processing and evaluation procedure, as presented by CEA [2009]. The Environmental Information System currently integrates about thirty databases for different purposes. There are mainly SQL databases and some of them are SQL/SDE spatial databases. For example, Figure 3 shows the application concerning water quality on the beaches of the Adriatic coast using data from the corresponding spatial database. All information about databases, technical specifications, their availability and potential users can be found in the Catalogue of Databases on the CEA web site, as shown by CEA [2009].

Figure 3. Example of the database for water quality on the Adriatic beaches available from CEA's Catalogue (*taken from CEA [2009]*)

According to the Environmental Agency plan to the year 2012, a total of seventy databases will be developed and connected into an integrated environmental information system. This environmental information system will be an open platform for interconnection with international environmental information systems like EEA/EIONET, GMS/INSPIRE, SIZE, WISE, etc., according to CEA [2009].

6. CONCLUSIONS

The most significant ICT based research activities of Croatia on several environmental projects at the national or international levels are presented and discussed. They concern investigations of the Adriatic Sea, water quality management in the Danube River Basin, and implementation of river information services. Since we have personally contributed to some of these activities, we have herein presented and discussed some interesting results.

Integrated environmental management and reporting, carried out by the Croatian Environment Agency, are described as well.

Results of the Adriatic Sea investigations concern the conceptual model of the referral information system on performed investigations, application of remote sensing methods and development of an automatic in situ measurement system. Development of sophisticated environmental information systems was first initiated within the Environmental Programme for the Danube River Basin (EPDRB) in which Croatia actively participated. For efficient management of environmental data, integration of GIS and the database for water quality was performed. The developed Web GIS application enables reporting on selected monitoring stations from the digital map of whole Danube River Basin. As a part of the CRORIS project, a specialized geographic information system for surveillance of the Croatian segment of the Danube River is being developed (*i.e.*, *River IS*). Several applications were developed using web and XML technologies to enable access and integration of data from different sources, timely presentations in the form of digital maps, diagrams and reports and remote control of measurement devices.

The national environmental information system was recently developed by the Croatian Environment Agency. In this way, integrated data management at the national level and reporting in line with the European requirements will be enabled. This is seen as an open platform for further interconnection with other European information systems.

REFERENCES

Bec, S., J. Pecar-Ilic, Z. Skocir, Alarm Reporting in River Information System, K. Tochterman, A. Scharl (Eds.), *Proceedings of the 20th Internatinal Conference on Informatics for Environmental Protection – EnviroInfo 2006*, Septtemer 6-8, Graz (AT) pp. 27-34, 2006.

CEA, Information about Environmental Information System, *Croatian Environment Agency Web site*, Zagreb, http://www.azo.hr/default.aspx?art=542, available on February 2009.

CRUP, General information on CRUP projects, *Inland Navigation Development Centre (CRUP) Web site*, Zagreb, http://www.crup.hr/eng/index.php, available on February 2009.

Dadic V., M. Bone, G. Beg Paklar, B. Grbec, D. Ivankovic, F. Matic and M. Morovic, Automatic meteo-ocean station (AMOS): real-time data acquisition, validation, archiving and numerical modeling, *Acta Adriatica* 47. (Suppl):133-148, 2006.

EEA, UNEP/MAP, *EEA UNEP/MAP Joint Work Plan 2006-2008*, MAP special publication, 2007

European Commission, Directive 2005/44/EC of the European Parliament and of the Council of Sep-07, 2005 on harmonised river information services (RIS) on inland waterways in the Community, *EC Official Journal L 255, 30/09/2005*, pp. 0152 – 0159, 2005.

GIS Forum Danube, Data Warehouse for the Danube Waterway (D4D) Brochures, *GIS Forum Danube*, 2004.

Lack, T., The Work of the Monitoring, Laboratory and Information Management SubGroup (MLIM/SG), in I. Lyons Murphy (Ed.), *Protecting Danube River Resources*, Kluwer Acad. Publ., The Netherland, pp. 127-134, 1997.

Pecar-Ilic, J., I. Ruzic, I., Z. Skocir, Information System on Investigation of the Adriatic Sea, *Croatica Chemica Acta*, Special issue dedicated to MARINE CHEMISTRY, Croatian Chemical Soc., 70 (1), pp. 435-458, 1997.

Pecar-Ilic, J., Information System for Temporal and Spacial Presentation of Complex Data, *PhD Thesis*, Center for Marine and Environmental Research, Zagreb (HR), 2001.

Pecar-Ilic, J., I. Ruzic, "Dynamic Web Application for Temporal and Spatial Presentation of Environmental Data", Sustainability in the Information Society, *Proceedings of EnviroInfo 2001*, Zuerich, L. M. Hilty, P. W. Gilgen (eds.), Published by Metropolis Verlag, Marburg, Germany, Part 1, pp. 431-436, 2001.

Pecar-Ilic, J., Z. Skocir, I. Ruzic, The Concept of Dynamic Temporal and Spatial Data Presentations on the Web, *Proceeding of the International Conference on Telecommunications*, Volume 1, Beijing, China, pp. 718-722, 2002.

Pecar-Ilic, J., I. Ruzic, Sharing Data among Users through Dynamic Web Applications for the River Information System, In: Minier, P., Susini, A. (Eds.), *EnviroInfo 2004*, Part II, CERN, Geneva, pp. 66-74, 2004.

Pecar-Ilic, J., I. Ruzic, Application of GIS and Web technologies for Danube waterway data management in Croatia, *Environmental Modelling & Software*, Special issue *Environmental Informatics*, ELSEVIER, 21 (11), pp. 1562-1571, 2006.

Pfliegl, R. (et al.), Introduction of EGNOS in River Information Services, *Proceedings of the 46th International Symposium on Electronics in Marine (ELMAR-2004)*, Zadar, pp. 137-142, 2004.

Project Jadran, Information about Croatian National Monitoring Programme Jadran, *Jadran project Web site*, http://www.izor.hr/jadran/, available on February 2009.

Ruzic, I., J. Pecar-Ilic, N. Hak, World Wide Web Pages for the Trans-National Monitoring Network of the Danube River Basin in Croatia, *Proceedings of International Conference Management of Transboundary Waters in Europe*, pp. 441-446, Mrzezino (PL), September22-25, 1997

Ruzic, I., J. Pecar-Ilic, Integration of GIS and Relational Database for the Trans-national Monitoring Network in the Danube River Basin, *Proceedings of International Workshop on Information for Sustainable Water Management - MTM III*, Nunspeet (NL), pp. 457-465, September 25-27, 2000.

UNEP, *Monitoring Programme of the Eastern Adriatic Coastal Area*, Report for 1983-1991, MAP Report Series No. 86, Athens, 1994.

UNDP/GEF, ISEP: Development of the Danube Information System for the International Commission for the Protection of the Danube River (ICPDR), *UNDP/GEF Final Report*, 74 pp. 1999.

UNDP/GEF, Geografic Information System for the Danube River Basin, *UNDP/GEF Danube Regional Project*, Final Report, 2006.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Information Science and Environmental Information Systems in Sweden

Åke Sivertun

FHS, Stockholm, Sweden (Ake.sivertun@fhs.se)

Abstract: Sweden as other countries are involved in the information age with transformation of data and information into databases and information systems. This is also valid for spatial information as maps and other location based data. A great part of environmental data are possible to give a place, an area or an extension and are dynamic so that they can be described and indexed in Geographical Information Systems. There is often an advantage to describe and model environmental factors together as to analyze sources and targets for environmental change. There is now a slow change in paradigm towards new ways to analyze data with such a spatial focus as it gives the opportunity to in detail monitor where the impact is bigger than the carrying capacity in nature and the impact on populations at risk.

In Europe we have the INSPIRE initiative with the goal to ensure establishment of a European spatial infrastructure. This initiative is also working in Sweden with focus on standardization of data and building meta data bases for easy access to data. However there are several aspects missing in the attempts to build such infrastructure. The business model is mainly based on charges for data – something that is rarely successful when an emerging technology has to be formed. There are also limited founds for research and development directed towards the intelligent use of these new information recourses. Building a spatial data infrastructure is rather expensive and the economic benefits and usability depend on a wide use of the information resources. Such wide use is further necessary to be able to collect, monitor, analyze and suggest counter measures to the great changes and challenges that will follow from changing climate. Other areas that could benefit from coordinated research on building environmental spatial databases are the inventory of natural resources for sustainable utilization and protection of these resources from natural as well as man-made hazards.

By introduction of GIS based transport planning and navigation tools Japan estimate to have reduced the number of driving kilometres with 10% with accompanying lower consumption of fuel, less emissions, more efficient transport systems and less accidents. By allowing the car industry as well as the producers of GPS and navigation systems to use the road databases with a nominal fee the technology could be successfully introduces and the number of customers for updates could rise from perhaps thousands to millions. In Finland the Environmental protection Institute have published and allows WEB-access to maps in GIS format (ESRI shapes) to allow research, administrative use as well as for the public. In this way the number of users increased from 300 to 3000 within a year! In Sweden we still have a model where the users (also within the public sector) have to pay for the data. There are some programs where universities have access to some geographic databases but there have never been any national initiatives or program to promote research and development of GIS like in the UK, in Netherlands or in the USA. This shortcoming is now (according my opinion) hindering the effective development of location based services and products. This is really sad as Sweden and many of the other Nordic countries traditionally have a very good situation regarding digital maps and spatially indexed environmental data. In the Nordic countries we also have unique possibilities to compare environmental impact on health through our public health system, very good health registers and population data on individual level that allow us to calculate dose/exposure/response – like the health effects for several environmental factors like radon and cancer. In a European perspective Sweden and our neighbours cold be a test field for such research as to formulate policies and allow further studies on an European level.

Keywords: Geographical Information Systems; INSPIRE; Sweden.

Sensitivity and Data Uncertainty Analysis of an Environmental Distribution Model

J. Urbánek[a], **J. Komprda**[a] **and K. Kubošová**[a]

[a] *RECETOX, Masaryk University, Kamenice 126/3, Brno, Czech Republic*
(urbanek@iba.muni.cz)

Abstract: Persistent organic pollutants (POPs) make up a part of the chemicals which pollute the environment. They are dangerous to people and animals in very small amount due to their specific properties. That is why people want to predict the fate of these chemicals in the environment. In the paper we describe a box model simulating the fate of POPs in the environment. Box models represent a special part of the environmental distribution models and as all other models (i.e. not only the environmental ones) are influenced by three categories of so called uncertainties: (a) uncertainty in the science itself and algorithms of the model, (b) data uncertainty, and (c) uncertainty regarding the appropriate application of the model. The first type (a) describes the lack of science and is therefore the most complicated to estimate. We, therefore, focus on the others. Mathematical description of the processes in the model leads to the system of ordinary differential equations. Every single coefficient in this system is inexact with some variability. The equations contain parameters whose values have to be measured (for example concentrations, fugacity capacities …) and thus are influenced by measurement errors; on the other hand they contain several simplifications such as taking the long-term means of the unsteady variables, for example the mean of temperature or the mean of the rain density. From the sensitivity analysis we can reach deeper understanding of the output uncertainties. We focused mainly on the sensitivity analysis of the previously mentioned variables: temperature and the rain density because they were greatly related to several parameters (like partition, sorption and leaching flux coefficients, concentrations …). For example the concentration of chemical in the air can change dramatically in a matter of minutes during raining because the POPs are very liable to a wet deposition. The results show significant changes in solution which confirm the expectations that the concentrations of chemicals (and the model itself) are strongly dependent on the mentioned parameters. By expressing all the uncertainties together with the sensitivity analysis we are able to tell how closely the model approximates the reality and what the key parts of the model which should be as precise as possible are.

Keywords: Modelling, Box model, Uncertainties, Sensitivity, ICT tools

1. INTRODUCTION

Persistent organic pollutants (POPs) are chemical substances that persist in the environment, have toxic properties, bioaccumulate, and are liable to long range transport. They rise due to the incomplete combustion of organic matter or fossil fuels. POPs are dangerous to people and animals in very small amount because of their specific properties. It is important to stop their production and try to prevent their rise. Human health and the environmental protection from POPs lead to an agreement on so called Stockholm convention in 2001. The convention came into force in 2004 and requires Parties to take measures to eliminate or reduce the release of POPs into the environment. An essential part of the control over such chemicals belongs to the understanding and predicting their fate.

This paper presents the sensitivity and uncertainty analysis of the Mackay-type multimedia model, so called box model [Mackay, 1991], [Cowan, 1994], which we have used for predicting the fate of POPs in the Czech Republic [Komprda, 2009].

2. THE MODEL

The model consists of three compartments: air, vegetation and soil. The soil compartment is further divided into seven horizontal layers for a better simulation of its heterogeneity. Moreover, each of these layers has two fractions (basic and deep). Altogether, there are 16 parts among which transport processes (e.g. diffusion, deposition, advection, degradation etc.) of pollutants occur. By expressing the mass balance equations for the model we obtain a system of 16 ordinary differential equations (ODEs).

We suppose there are no transport processes between our model and surroundings except for leaching from the bottom soil layer. Furthermore, we assume that there is no production (or rise) of monitored chemicals. Measured pollutant concentrations represent initial values of the ODE system. Concentration progressions of chemicals in time in each compartment are obtained as an output. For the purposes of model testing the concentrations in the air were set to be constant.

Typical POPs representatives are chemicals called polycyclic aromatic hydrocarbons (PAHs). We chose three of them (acenaphtylene, benzo(a)pyrene and pyrene) for model testing to include all their characteristics. The model containing 39 parameters is influenced by uncertainties like others. This study focuses on data uncertainty. So far we have used parameter values presented in previous research or their long term means (from the Czech Hydrometeorological Institute - CHMI). Several parameters like temperature or rain density are varied; values of others are rather uncertain. We performed sensitivity and uncertainty analysis of the model to find out which parameters mostly influence the output and what should be the next step in improving the model.

3. SENSITIVITY ANALYSIS

The aim of sensitivity analysis is to determine how sensitive the output is to changes in the input. Sensitivity is defined as the relative effect of a parameter on the result [Beyer, 2001]. There are two main techniques of sensitivity analysis: local and global one. Local sensitivity analysis estimates the effect of the variation of a single factor keeping all the others fixed at their nominal values. On the other hand global sensitivity analysis estimates the effect on the output of a factor when all the others are varying, enabling the identification of interactions in models [Cariboni, 2007].

We can easily rewrite the definition of local sensitivity in a mathematical (normalized) form using derivatives:

$$S(X_i) = \frac{\partial Y}{\partial X_i} \cdot \frac{X_i}{Y}, \tag{1}$$

where Y is the output value and X_i is the i-th input factor. We are not usually able to express the derivatives of the model output analytically, so we use numerical methods. The simplest and often used one is a Finite difference approximation:

$$S(X_i) \approx \frac{\Delta Y}{\Delta X_i} \cdot \frac{X_i}{Y}. \tag{2}$$

In the study we applied this method based on [Beyer, 2001], where the finite variation of parameter X_i was 10% (both up and down). Then the sensitivity can be expressed as:

$$S(X_i) \approx \frac{mean\ (\Delta Y_i^-, \Delta Y_i^+)}{Y} \cdot 10\ , \qquad (3)$$

where the minus sign in exponent represents a deduction of 10% (from a nominal value of X_i) and the plus sign adding same amount. According to Beyer, we can use the following classification scheme:

- low/no sensitivity: $S < 0.1$
- moderate sensitivity: $0.1 \leq S < 0.5$
- high sensitivity: $S \geq 0.5$

We carried out the sensitivity analysis for all three listed chemicals. Sensitivities were computed for pollutant concentrations in the vegetation and soil compartment. As for the soil phase we created the total soil concentration consisting of all the concentrations in the soil compartment with each concentration taking exactly that fraction which it gives to the total in real. Since there is no transport process between our model and surroundings except for leaching from the bottom soil layer, no production (or rise) of pollutant and constant concentrations in the air, the solution of the system leads to stable state. Values representing these stable states were taken into account while computing the sensitivities (see Table 1).

Temperature quantity can cause some problems while expressing its related uncertainty and sensitivity. 10% of the temperature nominal value (281.15 K) is 28 degrees which is inappropriate. We therefore took 10% of its value in centigrade degrees. Moreover, we assumed there are no transport processes requiring water within the phases in our model when the temperature is below zero.

Table 1. Most sensitive parameters for acenaphtylene (ACY), benzo(a)pyrene (BaP), and pyrene (PYR). Values higher than 0.5 are in bold.

Parameters and their units	ACY	BaP	PYR
	vegetation, soil	vegetation, soil	vegetation, soil
Partition coefficient air-water (--)	0.02, **0.87**	**0.96**, **1.17**	0.06, **1.11**
Partition coefficient octanol-water (--)	0.02, **0.56**	**0.95**, **1.13**	0.06, **0.62**
Soil density (kgm^{-3})	0.00, **1.01**	0.00, **1.01**	0.00, **1.01**
Specific biomass (kgm^{-2})	**0.99**, 0.01	**1.00**, 0.00	**1.01**, 0.00
Temperature (K)	0.04, 0.29	**0.89**, **0.97**	0.08, 0.45
Volume fraction of particles in air (--)	0.00, 0.12	**0.94**, **0.98**	0.06, 0.32
Aerosol density (kgm^{-3})	0.00, 0.12	**0.94**, **0.98**	0.06, 0.32
Leaf area index (m^{2}m^{-2})	**0.98**, 0.26	0.05, 0.00	**0.94**, 0.02
Mass transfer coefficient above vegetation (mh^{-1})	**0.98**, 0.24	0.05, 0.00	**0.94**, 0.01
Half-time of vegetation degradation (h)	**0.88**, 0.22	**0.90**, 0.00	**0.90**, 0.01
Chemical half-life in soil (h)	0.00, 0.35	0.00, **0.79**	0.00, **0.53**
Mass transfer coefficient of dry deposition to vegetation (mh^{-1})	0.00, 0.01	**0.76**, 0.00	0.05, 0.00
Rain rate (mh^{-1})	0.00, 0.35	0.19, **0.66**	0.01, **0.68**
Water fraction in soil (--)	0.00, **0.65**	0.00, 0.15	0.00, 0.44
Solids fraction in soil (--)	0.00, **0.65**	0.00, 0.15	0.00, 0.44
Scavenging ratio (--)	0.00, 0.07	0.19, **0.62**	0.01, 0.20

Figure 1 gives an illustration of the output changes for the most sensitive parameter both in vegetation and soil. Time scale was adapted to show the whole progression towards the stable state.

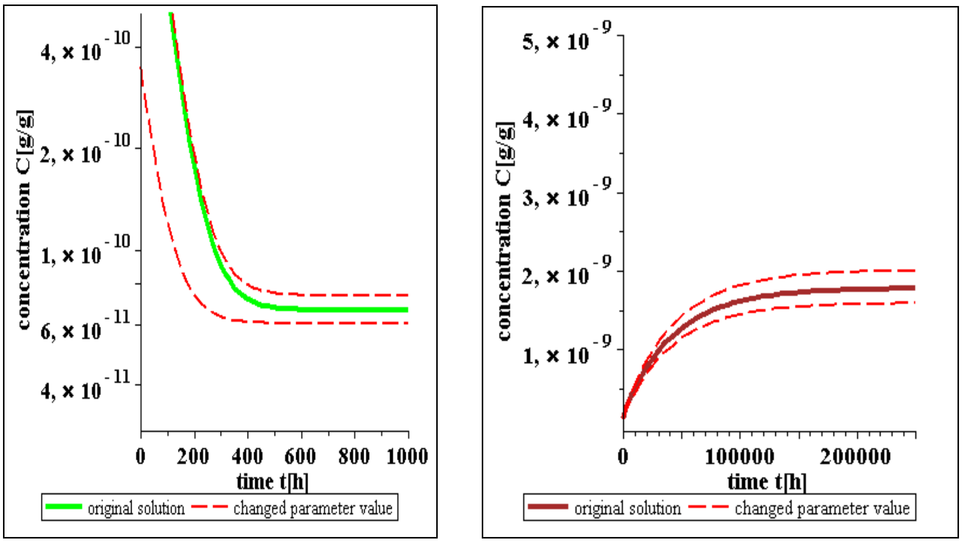


Figure 1. Concentration progressions for benzo(a)pyrene in vegetation (left) and soil (right) – original solution and solutions with 10% change in the partition coefficient air-water.

4. UNCERTAINTY ANALYSIS

Data uncertainty analysis was the next step in finding out the parameter influences on the model output. We assigned a probability distribution function to each sensitive parameter which we had the knowledge of according to previous research or our data (from CHMI). This was not feasible for all of them, so several parameters had to be assumed constant (those not listed in Table 2).

Table 2. Uncertainty distributions for model parameters

Parameters	Distribution	Source
Partition coefficient air-water at 298.15 K (acenaphtylene)	U($3.4 x 10^{-3}$, $5.9 x 10^{-3}$)	assumed
Partition coefficient air-water at 298.15 K (benzo(a)pyrene)	U($1.8 x 10^{-5}$, $3.0 x 10^{-5}$)	assumed
Partition coefficient air-water at 298.15 K (pyrene)	U($4.4 x 10^{-4}$, $7.6 x 10^{-4}$)	assumed
Partition coefficient octanol-water at 298.15 K (acenaphtylene)	U($3.6 x 10^{3}$, $1.2 x 10^{4}$)	assumed
Partition coefficient octanol-water at 298.15 K (benzo(a)pyrene)	U($7.1 x 10^{5}$, $2.7 x 10^{6}$)	assumed
Partition coefficient octanol-water at 298.15 K (pyrene)	U($5.9 x 10^{4}$, $3.3 x 10^{5}$)	assumed
Soil density	T(1000, 2000, 3000)	assumed
Specific biomass	U(1.0, 10.0)	assumed
Temperature	C(255.2, 300.4)	own value
Volume fraction of particles in air	L($2.0 x 10^{-11}$, 0.2)	[Luo]
Aerosol density	L(2400, 0.2)	[MacLeod]
Leaf area index	T(1.6, 3.0, 6.1)	assumed

Parameters	Distribution	Source
Rain rate	C(0.0, 2.1x10^{-3})	own value
Chemical half-life in soil (acenaphtylene)	L(1340, 0.6)	assumed
Chemical half-life in soil (benzo(a)pyrene)	L(5472, 0.6)	assumed
Chemical half-life in soil (pyrene)	L(13360, 0.6)	assumed
Water fraction in soil	U(0.2, 0.3)	[Beyer]
Solids fraction in soil	T(0.0, 0.6, 0.7)	[Beyer]
Scavenging ratio	U(2.0x10^4, 2.0x10^6)	assumed

Distributions: U = uniform (minimum and maximum values between parenthesis), T = triangular (minimum, most likely, and maximum values between parentheses), C = Custom (minimum and maximum values between parenthesis), L = lognormal distribution (mean and standard deviation between parentheses). [Luo] = [Luo, 2007], [MacLeod] = [MacLeod, 2002], [Beyer] = [Beyer, 2001].

We used daily values (measured in years 1996 – 2006) from CHMI for temperature and rain density. The amount of them was 4018 values. With the Monte Carlo simulation we obtained the same number of samples for other parameters. Model output was evaluated for all the samples. Then the contribution factors (of each parameter from Table 2) in total output variance were determined from Spearman rank correlation analysis between model inputs and output. By assigning a rank to each value in the model input and output, the Spearman's method calculates the sums of the squares of the differences in paired ranks:

$$r_i = 1 - \frac{6\sum_{j=1}^{N} d_j^2}{N^3 - N}, \tag{4}$$

where r_i is the rank correlation coefficient between the model output and the i-th input factor, d's are differences between ranks of corresponding model input and output, and N is the sample size in the Monte Carlo simulation (in our case $N = 4018$). The contribution factor of the i-th input to output variance can be determined as [Luo, 2007]:

$$k_i = \frac{r_i^2}{\sum_i r_i^2}. \tag{5}$$

Table 3 shows contribution factors higher or equal to 0.1. Figure 2 below this table illustrates output variance (according to parameter distributions in Table 2) for benzo(a)pyrene chemical.

Table 3. Most contributing parameters to output variance for acenaphtylene (ACY), benzo(a)pyrene (BaP), and pyrene (PYR)

Parameters	ACY	BaP	PYR
	vegetation, soil	vegetation, soil	vegetation, soil
Partition coefficient air-water	0.04, 0.10	0.00, 0.00	0.00, 0.05
Partition coefficient octanol-water	0.14, 0.34	0.00, 0.03	0.00, 0.21
Soil density (kgm^{-3})	0.00, 0.17	0.00, 0.04	0.00, 0.09
Specific biomass (kgm^{-2})	0.00, 0.00	0.42, 0.00	1.00, 0.00
Temperature (K)	0.80, 0.04	0.57, 0.64	0.00, 0.33
Chemical half-life in soil (h)	0.00, 0.00	0.00, 0.22	0.00, 0.03
Rain rate (mh^{-1})	0.01, 0.30	0.01, 0.03	0.00, 0.28

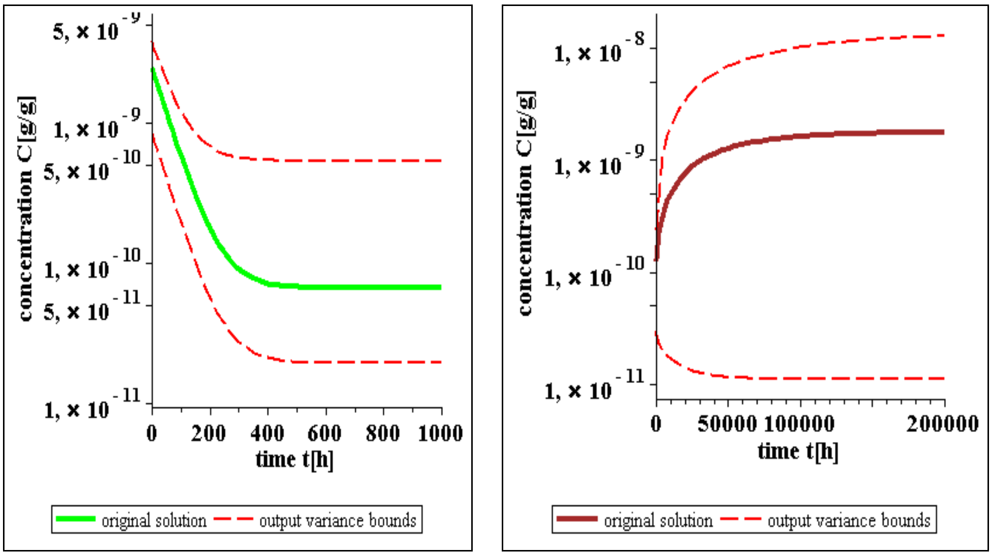


Figure 2. Concentration progressions for benzo(a)pyrene in vegetation (left) and soil (right) – original solution and the "highest and lowest" solution expressing the bounds of output variance.

5. DISCUSSION

The environmental distribution models are not the only ones which are influenced by three categories of uncertainties: (a) uncertainty in the science itself and algorithms of the model, (b) data uncertainty, and (c) uncertainty regarding the appropriate application of the model. It is very important to validate models and to conduct uncertainty analysis. Otherwise we cannot rely on the model results even when comparing with real observations. This paper presented data uncertainty as a first step of validating the model used to predict the fate of POPs in the Czech Republic.

Model parameters are usually constant values which approximate corresponding quantities. It is often therefore useful to include local sensitivity analysis as well to obtain effects of small parameter changes. One should keep in mind that local sensitivity analysis does not incorporate parameter interactions and that parameter sensitivity can be completely different in another setting. For this reason it is suitable to involve even global sensitivity analysis to get deeper understanding of data uncertainties.

We performed both sensitivity and data uncertainty analysis of the model for three different compounds to cover all their characteristics. The results show significant changes in solution confirming the expectations that the concentrations of chemicals (and the model itself) are strongly dependent on temperature. Other parameters like precipitation, partition coefficient octanol-water and specific biomass also highly influence the model output, but as we can see their sensitivities and contribution factors depend on the chemical type. Another important thing is that high parameter sensitivity does not have to mean high contribution factor.

For a proper data uncertainty analysis we have to know parameter value distributions which is often a huge problem. Moreover, Spearman rank correlation analysis assumes monotonic relationship between input and output. All the computations were done in the computer algebra system Maple. Using software we obtain another uncertainty due to finite number representation and applied numerical methods. Determination of such uncertainties, searching for parameter value distributions and global sensitivity analysis will be the next step in validating this model.

ACKNOWLEDGEMENTS

The work has been supported by the project INCHEMBIOL (MSM0021622412). The author wishes to thank Professor Jiří Hřebíček for supervising this work.

REFERENCES

Beyer, A., and Matthies, M., *Criteria for Atmospheric Long-range Transport potential and Persistence of Pesticides and Industrial Chemicals*, University of Osnabrück, 2001.

Cariboni, J., Gateli, D., Liska, R., and Salteli, A., The Role of sensitivity analysis in ecological modelling, *Ecological Modelling*, 203, 167 – 182, 2007.

Cowan, Christina E., *The Multi-Media Fate Model: A Vital Tool for Predicting the Fate of Chemicals*, SETAC Press, 1995.

Komprda, J., Kubošová, K., Dvorská, A., Scheringer, M., Klánová, J., and Holoubek, I., Application of an unsteady state environmental distribution model to a decadal time series of PAH concentrations in the Central Europe, *Journal of Environmental Monitoring*, 11, 269 – 276, 2009.

Luo, Y., and Yang., X., A multimedia environmental model of chemical distribution: Fate, transport, and uncertainty analysis, *Chemosphere*, 66 (8), 1396 – 1407, 2007.

Mackay, D., Multimedia Environmental Models: *Fugacity approach*, Lewis Publishers, 1991.

MacLeod, M., Fraser, A. J., and Mackay, D., Evaluating and expressing the propagation uncertainty in chemical fate and bioaccumulation models, *Environmental Toxicology and Chemistry*, 21 (4), 700 – 709, 2001.

Central and Eastern European Centre for Persistent Organic Pollutants (CEEPOPsCTR) – RECETOX, Masaryk University: http://www.recetox.muni.cz/ceepopsctr/index-en.php.

Maple: http://www.maplesoft.com

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

HYDROSYS - first approaches towards on-site monitoring and management with handhelds

Eduardo Veas, Ernst Kruijff, Erick Mendez

Institute for Computer Graphics and Vision
Graz University of Technology
(veas@icg.tugraz.at)

Abstract: HYDROSYS is a project targeted at improving monitoring and understanding of environmental processes and their management. The project introduces innovative concepts of on-site monitoring and event-driven campaigns. It builds an infrastructure that promotes synergy across user groups both on-site and at the workplace. This publication introduces initial steps towards defining on-site monitoring and an infrastructure to support it.

Keywords: Environmental information; sensor networks, handheld devices, spatial analysis.

1. INTRODUCTION

HYDROSYS is geared towards analysing environmental processes where they truly happen, in the field. Currently, there is a large gap between the work done in the field and that done in the office when analysing environmental processes. Environmental monitoring can be defined as the process of observing continuously and measuring regularly environmental parameters of a specific area in order to identify environmental changes and aid the decision making process related to the site. In that context, *on site environmental monitoring* comprises all activities that are done in the field such as identifying key variables of the problem, taking measures, data and images communication, understanding the site as a whole, and validation of a technical solution in order to optimize the management of natural areas. At all times, it is important to understand that on-site monitoring is not replacing nor can it be replaced by remote monitoring: it is a complementary task.

By advancing the state-of-the-art in on-site monitoring and management tools, the project aims at achieving greater synergy among users. A real situation will illustrate potential advantages of on-site monitoring and the HYDROSYS toolset. In summer 2008, cantonal administration received an extreme rainfall warning from MeteoSuisse for the Dranses catchment. They asked all cantonal natural hazards experts to stay at the headquarters in Sion and inform municipalities of this warning. During the event, cantonal experts analysed data coming from sensing stations. They asked through mobile or radio onsite observers to describe the situation, informed municipalities and tried to set up a coordinative effort with local municipalities. However, they were “blind” and could only consult the workplace data of the area. Cantonal experts asked the person on-site to give his or her opinion and to take some measures in order to quantify risks. After the natural disaster happened, cantonal experts asked the onsite observer to show them the event and to take some measures in order to quantify it. Cantonal experts in conjunction with local experts and municipalities sent guidelines and actions to the on-site observer to coordinate onsite actions.

Based on this real example, on site monitoring and HYDROSYS technology will develop communication and exchange methods (e.g. images, data, graphs) from an on-site observer to decision makers who are generally at workplace and vice versa. In fact, end users:

- will have a better view of the global situation before, during and after an event with the image transmission and overlay techniques.
- will save time and money doing onsite and online analysis

- will have a better visualization of the field when they are in office or headquarters

On-site monitoring tools do not replace any work carried out by the engineering companies but give additional information at a higher temporal resolution. It is important to state is that the decision-making process will not be automated in HYDROSYS, but happens "in the head" of the user, helped by sensor readings. This publication describes the basic technical framework, gives a short application example and introduces the envisioned visualization of environmental data in the field.

2. RELATED WORK

Currently, environmental monitoring is mostly done by an on-site observer, few expensive sensing stations with data logger (e. g. IFKIS network in Switzerland) and a Geographical Information System (GIS). With this current approach, many processes cannot be monitored. Sensor data is still (timely) limited and scattered in the environment: even when a problem is detected by a sensor, it can often not be traced back directly, since detailed information on the area is missing. As a result, many processes are badly understood and both their representation (including the physical process model) and their visualization are incomplete [Barrenetxea et al, 2008]. Moreover, current sensing stations have several limitations, such as no real time data and few spatial coverage, expensive cost, etc. In hydrology, for example, there have been only limited field campaigns with in-situ spatial observations [Barrenetxea et al, 2008]. In that context, being able to capture spatial and temporal variability of environmental parameters, to develop real-time image and data transmission [Shin, 2007] and data visualization in order to enhance environmental model, prediction tool and decision making process will be the next challenge in environmental monitoring [Bogue, 2008]. Therefore, current trends are moving toward deploying a large number of wireless sensing stations in order to provide high spatial and temporal density measurement. Wireless Sensor Networks (WSN) [Aberer et al, (2007), Chong & Kumar (2003)] have become a widespread tool for monitoring a wide range of environmental phenomena. Many research projects are investigating possible applications of sensor networks ranging from habitat monitoring [Szewcszyk et al, 2004] to agriculture [Langendoen et al, 2006] and to environmental monitoring [Selavo et al, 2007; Werner-Allen et al, 2005]. However, deploying a WSN in the field has always been reported as a difficult task and remains challenging [Barrenetxea et al, 2008].

3. APPLICATION SCENARIOS

Several application scenarios have been chosen as test cases for HYDROSYS. They are based on sites of interest in Switzerland (La Fouly, Gemsstock, Dorfberg) and Finland (Ridalinpuro and Kylmäoja) each emphasizing different aspects of the system.

La Fouly

This scenario aims at providing good quality data to model hydrological processes in steep alpine catchments. The data will feed a 3D, distributed and physically based hydrological model reproducing processes, such as infiltration and redistribution of soil moisture, water runoff, darcian subsurface flows, evapotranspiration, snow and ice melting. Results from distributed input will be compared to those obtained from a single meteorological station to investigate the relevance of geospatial variability in watershed modeling.

Gemsstock

In the Alps generally permafrost can occur at altitudes above 2400 m. Increasing temperatures in rock faces and ice filled fissures can have major impact on ground stability causing damage to infrastructure in these areas (such as cable car pylons and stations) affecting their usability. Deformations on buildings/pylons have to be surveyed accurately and regularly by an expert (e.g. engineering companies). The new tool would be applied on a daily/weekly basis to visualize onsite deformation rates and to receive feedback on the stability of the infrastructure. The onsite monitoring tool does not replace any work carried out by the engineering companies but would give additional information on deformations at a higher temporal resolution and could be carried out by the people working onsite.

Dorfberg

Wet snow avalanches are hazards of high importance showing up with a high frequency and a high degree of potential damage to infrastructure in mountainous regions. So far the processes which cause the formation and triggering of wet snow avalanches are poorly investigated. With steep slopes and its southwards aspect Dorfberg is a good place for doing wet snow avalanche research. In recent years many small wet snow avalanches, which occur on a quite regular frequency and also some big events have been monitored in the area.

Ridalinpuro

Ridalinpuro is a small area, which is currently being studied and constructed according to the principles of environmentally sensitive hydraulic engineering. There is good data from the site and it is (e.g., already equipped with v-notch weirs). The measurement campaigns on Ridalinpuro site will investigate the impact of land use change on urban runoff and the effect and pollution abatement efficiency of the environmentally sensitive hydraulic structures that are being built on parts of the site.

Kylmäoja

Kylmäoja is a larger urban catchment, which is experiencing pressure from human activities. The hydrological and hydraulic data that is available from Kylmäoja is limited. The measurement campaigns on Kylmäoja will investigate the impact of land cover, construction, and other human activities on water quantity and quality. The construction of the logistic centre in the upper part of the catchment and the activities on the major airport that is partly on the catchment make Kylmäoja a suitable site for the study.

4. SYSTEM OVERVIEW

The following use case attempts to clear HYDROSYS tools can be utilized. It also provides some hints as to how the system components relate:

Silvia arrives at La Fouly, a site where she means to analyse water discharge. Upon arrival, she starts her “hydrology researcher” profile, which allows her to visualize continuos data from sensors. While going through the site, she often points her handheld at the environment, on the screen she sees a live video feed overlaid with precipitation levels from the day before and current moisture maps, taking notes. At some point, a message from Vincent tells her that some sensors are no longer in the appropriate position. She brings up her “sensor specialist” profile, which allows her to visualize sensors and their status, screening out all readings. She switches to a map view and finds Vincent’s location, when turning towards him, on handheld screen she can already see the failing sensor, and the displacement from the appropriate location, they set about to fix the sensor….

In the use case, a team of users is working on a site. The data about the site spans several activities and every user may access data that is important for them. This data can come from continuous sources or from static ones, processed previously. In any case there is a flow of data coming from these sources, being processed and then visualized by users. The following diagram shows the overall connectivity of high level components in HYDROSYS. Each of these components performs a specific task for the system, and they are not necessarily executed concurrently or in the same computer.

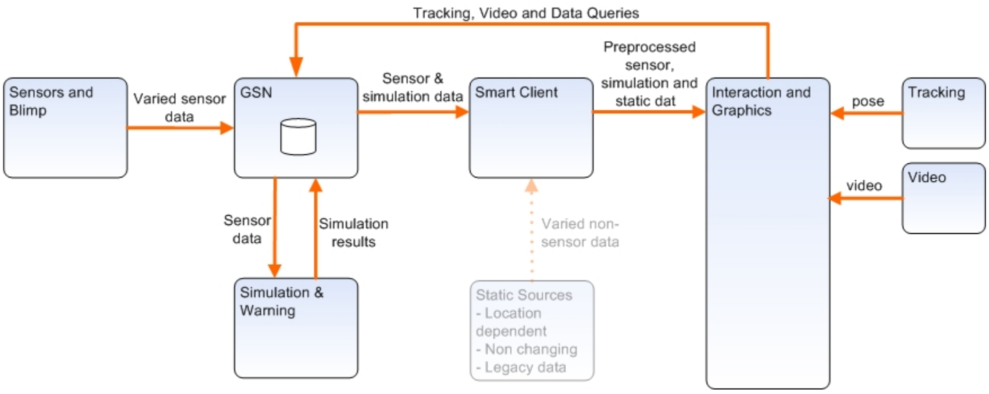


Figure 1. Software architecture diagram

Data is generated and pushed into the sensor network by the Sensors and Blimp component. Common sensors include sensorscope stations, temperature, moisture and water discharge sensors and a blimp that produces textured models of the environment.

Simulation and Warning: acts just as another sensor performing all the simulation tasks of the project. It reads sensor data from GSN and outputs results to it. GSN is a distributed sensor network that stores and propagates data during field work. Static Sources are a data source separate from GSN, providing mainly DEM/DTM and other legacy data. All these data sources form a rich information space that is accessed by HYDROSYS. The Smart Client component is a negotiator for data. It gathers information based on user needs and preferences (stored in a profile). It also prepares information in a form that can be efficiently handled by mobile devices, performing appropriate conversions (reducing data, preparing 3D models, filtering information) for each platform. Interaction and Graphics leverages advanced visualization techniques to render and overlay views of the data depending on the user's location. It enables interaction through visual controls and communication tools. Interaction and graphics needs to accurately know the pose (position and orientation) of the user in order to correctly overlay information. Tracking provides pose information to mobile devices. It is a hybrid solution supported on multiple devices and techniques. The visualization component also requires a view of the real world which is obtained from the Video component.

5. SENSORS

HYDROSYS creates a rich information space integrating diverse sensors and information sources. Multivariate sensors and sensor stations are available that potentially provide around 20 different kinds of environmental sensor readings. Environmental sensors are directly related to case study and location.

Beside environmental sensors, multiple cameras are part of a multi-camera framework for sensing and reconstruction purposes, including a camera supported on a pan-tilt unit for remote sensing. The latter can be remotely controlled by panning-tilting in the desired direction. User devices hold cameras that can also be considered part of the camera framework and used to share viewpoints on the location. Additionally, an unmanned aerial vehicle (blimp) can be used with two cameras, a normal range and a thermal camera. The blimp will provide a live, up to date surface model of the ground in the form of a height map, temperature map and colour map. The blimp will be flown prior to the initial experiment in order to generate a high precision height which will be optimized off-line. This model will be used during the live experiments. The model will then be refined in an off-line step, from the experimental data, making topographic data at different times available. The combination of these sensors generates highly detailed data for the site at hand, improving the performance for on-site monitoring tasks.

6. INTERACTION AND GRAPHICS

One advantage of on-site monitoring is the fact that the actual site, not some probably outdated model of it, is available to the user. A challenge for the tools developed by HYDROSYS is to exploit this fact, relating synthetic data to the real world and attempting to convey as much information as possible without loosing the real world context.

The technique used to merge real world views with synthetic data, known as augmented reality (AR), allows the user to walk around and observe the environment, continuously getting a "correct view" on the sensor data. The task of augmented reality is to render computer generated artefacts correctly registered with the real world in real time. The term "correctly registered" means that these artefacts must appear in the appropriate position relative to the point of view of the user. One can imagine two layers, one with a video background and another one where the system renders graphical information and which is overlaid on top of the first one. The term real time often means that these overlays must be generated at least at 15 frames per second to give the human eye a sense of coherence.

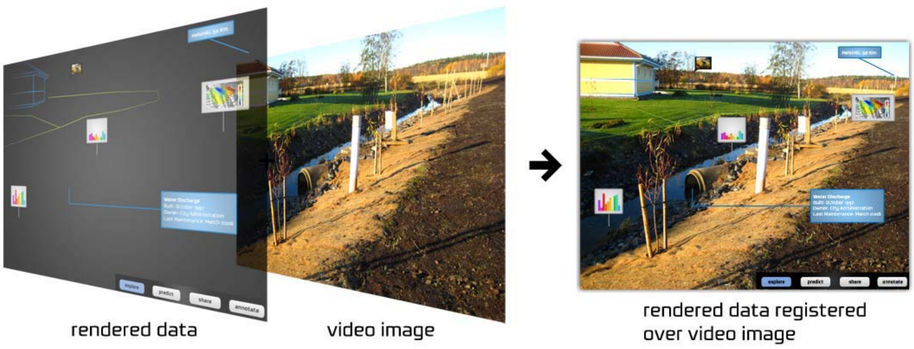


Figure 2. Example of augmented reality techniques to overlay data in the real world

Interaction and Graphics is responsible to present data to the users in a format that helps them better understand it and relate it to the real world, and to provide tools navigate, modify the data and behaviour of the system. The challenges in this area have to do with presenting graphics in small screens, avoiding visual clutter, conveying information correctly as well as proposing an intuitive interaction metaphor for handheld devices.

6.1. Visualization

SensorScope and other sensing stations measure key environmental data such as air temperature and humidity, surface temperature, incoming solar radiation, wind speed and direction, precipitation, soil moisture and water suction and so on. The measured information comes in a numerical form or even perhaps as voltage readings. The raw presentation of said numerical information is unsuitable as the cognitive load of understanding the readings is too high. Instead, hydrologists conventionally use mathematical software to display the abstract information delivered by the sensors on a more human readable form, often as plotted graphs. AR visualization requires several sets of 3D models: sensor readings in graphical form, digital terrain models (DTM), glyphs for known objects (such as other users in the field). The traditional pipeline for generating interactive graphics based on sensor data is the following: Given a set of sensor readings, the system generates polygons with corresponding values associated to them. These values may represent colors, hues, brightness, and other visual properties. These polygons are generated on the location given by the context data of the sensor. All sensor stations, and simulation results will provide data that is already geo-referenced and directly useable for the generation of visualizations.

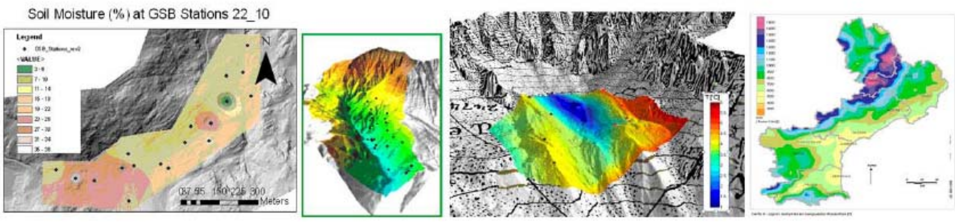


Figure 3. Example of heterogeneous data and visualization forms.

Sensor data can have different visualization forms, that all have their specific advantages and disadvantages to interpret the data. The current way we intend to interact with sensor data layers is organized by browsing through the different visualization modes.

Focus + Context

When visualizing large amounts of information scattered around a site, it is sometimes important for the system to direct the attention of the user to possible areas of interest. There exist several techniques to achieve this such as distorting the image, overlaying conspicuous artefacts or modifying the visual properties of the scene. Of particular interest are those that modify the image's saliency, this is, the property of certain parts of an image to attract the user's attention more than others. During the course of the project, several techniques will be applied and experimented with including filtering, depth perception, explosions, and others.

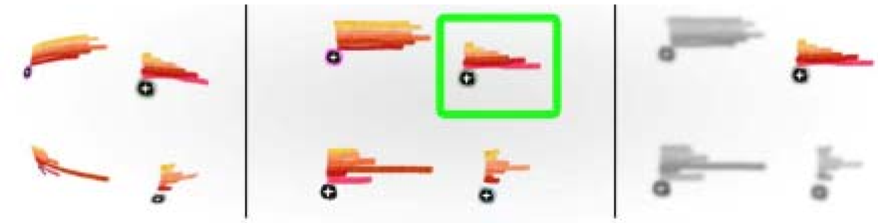


Figure 4. Directing the attention to particular sensor readings in the scene, by distortion, overlay and saliency modification.

6.2. Interaction

Interaction covers several issues ranging from the hardware devices to the user interface. The tasks related to hardware devices insure the platform is appropriate for AR. While those related to software user interfaces introduce means to operate and modify the behaviour of the system.

Handheld display system construction

The UMPC is one of HYDROSYS target lightweight devices. It is a fully featured personal computer for mobile use but still it has some strict limitations for usage as AR platform. UMPCs often need to be extended with sensors and controllers even when the UMPCs already contain such sensors. It is a matter of the quality required by AR applications that is addressed by additional sensors (high framerate and high resolution camera, high resolution and low error GPS, gyroscopes). A device aimed at AR applications enabling hardware extensions in an ergonomic, usable manner has been presented in [Veas and Kruijff 2008]. A new version of this device was tested with end-users in Finland (see Figure 4).

Figure 5. New construction used by Finnish end-users in the field (Nummela site)

User Interfaces for environmental monitoring with handhelds

In order to accommodate the front-end for the different actions, HYDROSYS will provide a componentized user interface kit. A higher level user interface (UI) toolkit (the widget toolkit) and a set of navigation techniques form the core upon which task-specific interfaces are based.

- AR user interface toolkit: The user interface toolkit consists of the widget toolkit, which offers higher level user interface elements like layers, forms, buttons, sliders, and so on. The widget toolkit focuses on data organization and visualization adaptation for complex spatial data sets to accommodate for the different hardware platforms. Information is organized in layers, for example accommodating a layer for each different type of data. Multiple layers are stacked on top of each other with the video image and DEM at the bottom. The user performs specific actions on layers to improve the visual quality of the overall interface, for example to avoid visual clutter. Such actions include blending layers in or out, or performing modifications on the rendered data or video image to advance readability. Reorganizing layers up or down the stack allows more important data to be rendered on top of contextual data, aiding the users in visualizing directly what they are most interested in. These techniques help users to customize their viewpoint on data. This has considerable perceptual and cognitive advantages. Certain techniques can potentially be related to the focus+context techniques being developed to improve the perceptual quality of the visualization

- Interfaces: these are extension modules aimed at specific tasks including sensor placement, simulations, collaboration. The sensor placement interface module supports planning and setting up sensor networks. The data produced by the sensors can be accessed through the data retrieval module that basically is the data selection front-end to the global sensor network (GSN). Users can perform simulations by using a simulation module to select data sets and simulation scripts. Finally, users can communicate, annotate and share data using collaboration tools. Most of these modules are integrated in the base interface, the interface used at the workplace, extended with functionality that enables users to plan a site and setup a campaign, and to perform post onsite analysis.

.

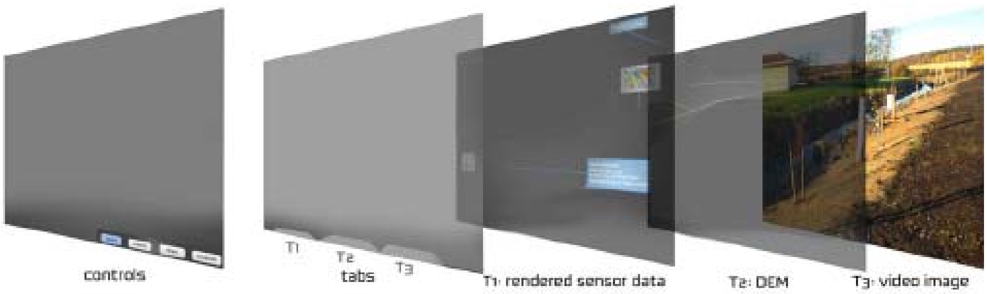


Figure 6. Different layers used in the AR visualization

Navigation and Viewpoint Change

The main factor exploited by Augmented Reality is that the application relies on the user's sense of orientation for visual understanding of the data presented. Navigation techniques associated with changes in viewpoint are included as aids for exploration of complex spatial datasets. Changing viewpoints using external cameras is a technique enabled by the characteristics of HYDROSYS deployment, having several cameras in the field, observing the location of interest from different viewpoints. At this point, it should be stated that one of the cameras being deployed is a pan-tilt unit. Using a simple front-end, users will be able to control the orientation of this camera. Similarly, users will be able to select a "hovering" spot for the blimp, to get a specific exocentric viewpoint of the site.

5. CONCLUSIONS

Aided by modern sensing technologies HYDROSYS extends current practice by enabling on-site data collection and visualization and supporting multidisciplinary teams of users in the sharing of their results. The project also provides a tool with which short-term events can be studied with appropriate detail.

Hydrosys leverages and advances state-of-the-art technologies in different disciplines from highly heterogeneous sensor networks, outdoor hybrid tracking, mobile augmented reality, image processing, down to the test case scenarios in alpine and nordic locations. The challenges that the project faces are manyfold, integration of these technologies being among the most crucial ones. Validation of proposed visualization and interaction techniques follows user studies to determine cognitive load of different techniques, the level of understanding achieved, intuitiveness of interaction and several other perceptual factors. These human factors, which are present in all systems, are stressed by the extreme conditions presented by HYDROSYS (small screens, large heterogeneous datasets, uncomfortable controllers, etc.). HYDROSYS will address these and other challenges at appropriate levels to insure the innovative advantages introduced by on-site monitoring can be achieved.

ACKNOWLEDGEMENTS

This work is partially funded through the EC-funded 7th Framework project HYDROSYS (224416, DG-INFSO) and is a cooperation between the following institutions and companies: Graz University of Technology (Austria), Ecole Polytechnique fédérale de Lausanne (Switzerland), Swiss Federal Institute for Forest, Snow and Landscape Research (Switzerland), Helsinki University of Technology (Finland), University of Cambridge (United Kingdom), Luode Consulting Oy (Finland) and Ubisense LTD (United Kingdom). Thanks to these partners for their contributions to this article.

REFERENCES

Aberer. K., G. Alonso, G. Barrenetxea, J. Beutel, J. Bovay, H. Dubois-Ferrière, D. Kossmann, M. Parlange, L. Thiele, and M. Vetterli. Infrastructures for a Smart Earth - The Swiss NCCR-MICS initiative, PIK - Praxis der Informationsverarbeitung und Kommunikation, 2007

Barrenetxea. G., F. Ingelrest, Y. M. Lu and M. Vetterli. Assessing the Challenges of Environmental Signal Processing through the SensorScope Project. *The 33rd IEEE International Conference on Acoustics, Speech, and Signal Processing* (ICASSP 2008). Las Vegas, Nevada, USA, 30 March - 4 April 2008

Barrenetxea. G., F. Ingelrest, G. Schaefer, M. Vetterli, O. Couach and M. Parlange. SensorScope: Out-of-the-Box Environmental Monitoring. *The 7th International Conference on Information Processing in Sensor Networks (IPSN 2008)*. St. Louis, Missouri, USA, 22-24 April 2008.

Bogue, R. Environmental sensing: strategies, technologies and applications*. Sensor Review SENSOR REVIEW.* 28.4, pp 275-282, 2008.

Chong, C.Y. and Kumar, S.P., Sensor Networks: Evolution, Opportunities, and Challenges. *Proceedings of the IEEE,* vol. 91, no. 8, pp. 1247–1256, August 2003.

Langendoen. K., A. Baggio, and O. Visser. Murphy loves potatoes: Experiences from a pilot sensor network deployment in precision agriculture. *In Proceedings of the IEEE International Parallel and Distributed Processing Symposium* (IPDPS), Apr. 2006

Selavo. L., A. Wood, Q. Cao, T. Sookoor, H. Liu, A. Srinivasan, Y. Wu, W. Kang, J. Stankovic, D. Young, and J. Porter. LUSTER: Wireless sensor network for environmental research. *In Proceedings of the ACM International Conference on Embedded Networked Sensor Systems* (SenSys), Nov. 2007

Shin, I. Development of an internet-based water-level monitoring and measuring system using CCD camera. *Proceedings of SPIE--the international society for optical engineering* ICMIT, 2007.

Szewcszyk. R., A. Mainwaring, J. Polastre, J. Anderson, and D. Culler. Lessons from a sensor network expedition. *In Proceedings of the IEEE European Workshop on Wireless Sensor Networks and Applications* (EWSN), Jan. 2004.

Veas, E. and E. Kruijff. Vesp'R - design and evaluation of a handheld AR device. *In Proceedings of the 7th IEEE and ACM International Symposium on Mixed and Augmented Reality* (ISMAR'08). 2008

Werner-Allen. G., J. Johnson, M. Ruiz, M. Welsh, and J. Lees. Monitoring volcanic eruptions with a wireless sensor network. *In Proceedings of the IEEE European Workshop on Wireless Sensor Networks and Applications* (EWSN), Jan. 2005.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Networking Environmental Informatics in Europe with Respect to Sustainability

Kristina Voigt

Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Institute of Biomathematics and Biometry, 85764 Neuherberg, Germany (kvoigt@helmholtz-muenchen.de)

Abstract: It becomes extremely urgent to consider environmental protection as a whole and environmental chemicals as one important aspect in connection with sustainability and sustainable development. Environmental modelling, environmental information systems, environmental information play a major role in this respect. The disciplines of Environmental Informatics, Environmental Statistics (Environmetrics), Environmental Modelling and Software, Chemoinformatics, Chemometrics, Sustainable (Green) Chemistry as well as the discipline of Partial Order in Environmental Sciences and Chemistry are introduced. In this paper the focus lies on Environmental Informatics. An overview on the activities working on similar environmental and chemical topics will be given focussing on the sustainability (green) issues of the groups. The small interplay among the disciplines is outlined and future activities will be encouraged. Furthermore future research is touched like the intensification of work in the field of green metrics and green / sustainable modelling is briefly outlined.

Keywords: Environmental Informatics; Environmental Modelling; Sustainability, Sustainable Chemistry; Partial Order

1. INTRODUCTION

These days the terms environmental protection in connection with sustainability and sustainable development receive an increasing interest in society. Sustainability has become a key concern for consumers and industry. However, coming to understand its implications in the context of daily life can be difficult. It may be hard to come to terms with the idea that traditional activities present environmental risks and may lead to real emergencies, manifested differently depending on local conditions. Everyone will be asked to do more but the move towards sustainability may require the total transformation of current attitudes and actions, see Hepting [2007]. An important issue in the context of sustainability is the topic of sustainable chemistry.

A report by the OECD (Organization for Economic Cooperation and Development) [2002] describes the consumers in member countries as generally having a high level of activism with respect to green issues, but a relatively low willingness to pay. There seems to be a decline in trust of nearly all sources of environmental information and an increase in confusion about prioritization of environmental goals and which actions can be most beneficial. Disciplines which are described in this paper which are in the interdisciplinary of environmental protection on one side and computer sciences, mathematics, statistics on the other side play a vital role in supporting the knowledge about the status of the environment and sustainable topics as well as the prioritization of environmental goals and actions.

Several important disciplines and societies are introduced and briefly described in this paper. However, the interaction and collaboration among these activities has to be improved in the future.

This paper is an update of a presentation given at the iEMSs Conference (International Environmental Modelling and Software Society) in Barcelona in 2008 by Voigt [2008a].

2. ENVIRONMENTAL AND CHEMICAL DISCIPLINES

2. 1 Overview of Disciplines

The overview of theoretical disciplines focussing on environmental and chemical sustainability issues is far away from being complete. It just reflects the activities in which the author is involved in some respect.

The WebPages, background literature as well as some journals in which the fields are covered are listed in Table 1.

Table 1. Overview on Environmental and Chemical Disciplines, Societies

Name of Discipline	Society	URL	Background material Examples	Sec
Environmental Informatics	TC Informatics for Environ-mental Protection	http://www.iai.fzk.de/Fachgruppe/GI/welcome.eng.html/	Proceeding volumes of annual conferences, Umweltinformatik, see Page [1995], Environmental Informatics, see Avouris [1995]	2.2
Environmental Modelling and Software	iEMSs	*http://www.iemss.org/society/*	Journal Environmental Modelling & Software (Jakeman)	2.3
Environmetrics	TIES	*http://www.iemss.org/society/*	Encyclopedia of Environmetrics, see El-Shaarawi, [2002], Environmetrics (El-Shaarawi)	2.4
Partial Order in Environmental Science and Chemistry	Hasse Expert Group	-	Proceeding volumes of workshops Partial Order in Environmental Sciences and Chemistry, see Brüggemann [2006]	2.5
Chemoinformatics	No society defined yet	*http://www.iemss.org/society/*	Handbook of Chemoinformatics, see Gasteiger [2003]	2.6
Chemometrics	International Chemometrics Society?	*http://www.iemss.org/society/*	Chemometrics in Environmental Analysis, see Einax [2004]	2.7
Green Chemistry /Green Metrics	- GDCh WG Green Chemistry, - Centre for Green Che-mistry and Green Engineering	*http://www.gdch.de/strukturen/fg/nachhaltigechemie.htm* *http://www.greenchemistry.yale.edu/about_the_center/*	Anastas [1997, 2002]	2.8

2. 2 Environmental Informatics

40 years ago, both environmental science and computers were in the starting blocks of a world-changing evolution. Environmental Informatics (EI) originates from application needs in engineering, e.g. data acquisition, computer-based process control, operation of measurement stations, data transmission, pollution modelling, and image processing in remote sensing. At the same time, developments e.g. in landscape ecology, biosciences,

identification, model building, system science, and environmental policy enriched the structural understanding of scientists working in this interdisciplinary area. E. U. von Weizsäcker dated the emergence of Environmental Informatics back to the time when no specialist area and no identifier for these activities existed, see Hilty [1997]. Today, the overall objective of Environmental Informatics is to process data, analyse information, formation systems related to the environment in its broadest sense, while using methods, techniques, and tools of computer science. With this, it is hoped to contribute to environmental protection, to support risk management, and finally to approach a more sustainable future. Early efforts started in the mid-1980s when time seemed to be ripe for applications of information systems in the emerging field of environmental protection. The nucleus of the upcoming field of "computer application for environmental protection" was the 1st Symposium held at the Karlsruhe Research Center in 1986 and the foundation of the Technical Committee "Informatik im Umweltschutz" (Environmental Informatics). The new extended title of the Technical Committee is now **Environmental Informatics, Informatics for Environmental Protection, Sustainability and Risk Management**. Around 23 years later, 23 EnviroInfo conferences have been held with approximately 3.000 papers by 7.500 authors published in proceedings' volumes since the first event. From a German initiative the conference has largely expanded into the European field and is also addressing international experts. Due to the preference and expertise of the symposium organizer, each event had its own profile and specific subjects. The EnviroInfo 2009 will be held at the FHTW Berlin University of Applied Sciences in Berlin from September 9th – 11th. The sub-title of the conference is Environmental Informatics and Industrial Environmental Protection: Concepts, Methods and Tools. For further information see *http://www.enviroinfo2009.org/*.

The history and the development of the past EnviroInfo Conferences are outlined in Figure 1. Further information not only concerning the history but also with respect to the goals and future activities are given by Pillmann et al [2006a].

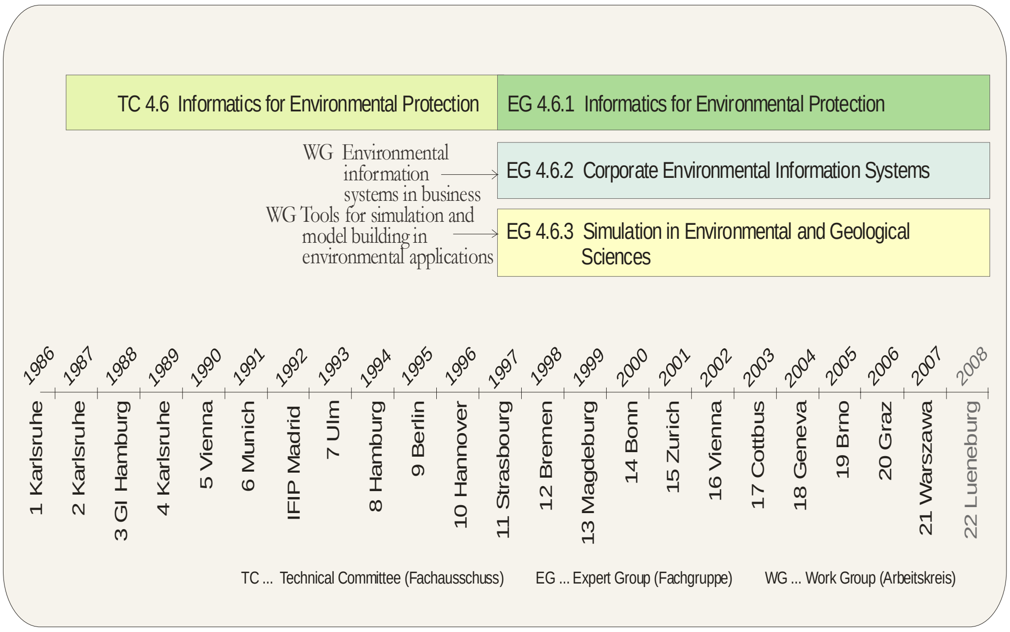


Figure 1. Development and Places of EnviroInfo Conferences 1986-2008 Pillmann [2006 a,b]

In order to understand the scientific field of Environmental Informatics it is useful to introduce an epistemological framework (a method to illustrate a scientific profile) for this discipline which was estabglished by Isenmann [2008]. Usually all research fields share a common "perspective", no matter of certain roots or specific origins. Put differently, any form of intellectual arrangement could be characterised through a specific "view". In terms of epistemology, such a "perspective/view" is understood as a specific way to treat and deal with "problems". These problems are phenomena that are regarded as being of relevance. For the clarification of Environmental Informatics' theoretical foundation with the help of tools of philosophy, here a generic framework is used. This framework is represented in the form of a pyramid and structured in layers. It serves as a rough "architecture" for examining epistemological issues when portraying the contours of a scientific profile along four basic layers (I-IV) or contexts respectively (Figure 2).

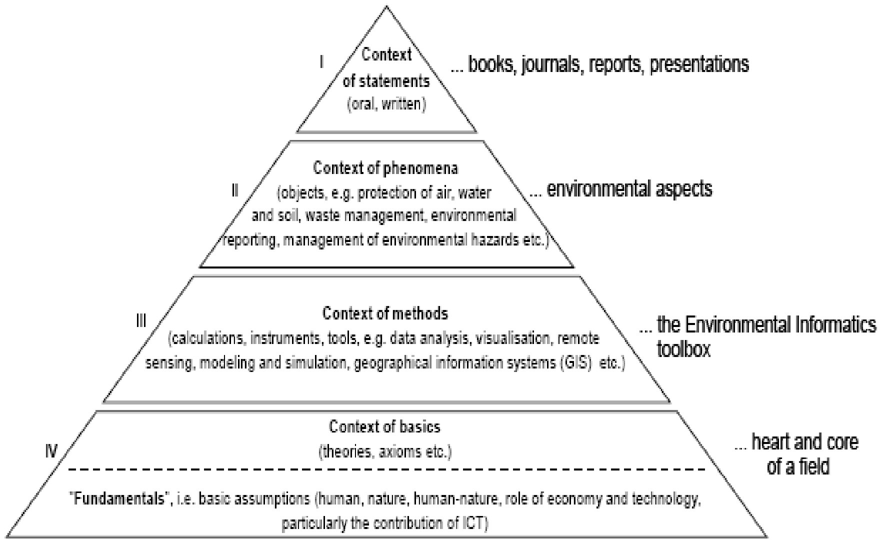


Figure 2. Scientific profile of Environmental Informatics, illustrated in a basic architecture, see Isenmann [2008]

The architecture for Environmental Informatics (Figure 2) has been conceptualised through certain issues that are widely discussed in the community, so far. Hence, these issues can be taken as prototypical for the current state of the field. For example, data analysis, visualisation, geographic information systems, environmental databases, modelling and simulation, and knowledge management are some of the methods that call the community its own and thus make the Environmental Informatics toolbox, among others by Page [1990], Page [1995], Rautenstrauch [2001]. The issues used to conceptualise the epistemological architecture have been identified from a document analysis of the EnviroInfo Conferences, and the current literature. The short list of issues again has been clustered in terms of the four layers.

A recent larger attempt to extend the focus from environment to sustainable development in general has already been established by the EnviroInfo community with the EnviroInfo conference in Zurich, Switzerland in 2001. The EnviroInfo Conference 2001 was sub-titled: Sustainability in the Information Society by Hilty [2001]. An approach to quantify the impact of information and communication technologies on sustainability was developed in a recent project by Hilty [2004, 2006a]. This is only a small sample of available background material concerning EI.

A paper on Environmental Informatics and Industrial Ecology, two emerging fields striving for sustainability provides an approach that could be used to compare the fields of research of Environmental Informatics and Industrial Ecology. Based on supposed overlapping areas, this approach further facilitates the linkage between their associated communities. From a theoretical perspective, the approach makes clear whether the fields of research may have common objects, similar tools, and shared principles, values, and value judgements. No less important, the approach is also helpful from a practical point of view as it helps to identify issues for fruitful institutional cooperation and joint projects. It is argued here that a proper method for a comparison between Environmental Informatics and

Industrial Ecology would be an epistemological point of view. Such an “out of the box” perspective finally helps to describe the emerging bodies of theory and clarifies contours of their certain scientific profiles. The proposed approach is illustrated in the form of a pyramid (see Figure 2), structured in layers like architecture, and built with the help of tools of philosophy. It has its methodological basis in a generic framework used in epistemology. This generic framework is further conceptualised through document analyses identifying certain issues that are prototypical for the current communities and literature both, for Environmental Informatics and Industrial Ecology by Isenmann [2008].
An important initiative of the Environmental Informatics Society is the ICT-ENSURE project. ICT-ENSURE (Information and Communication Technologies - Environmental **Sustainability** Research) is the leading support action in the area "ICT for Environmental Sustainability Research". The main objectives are: to extend the network of environmental sustainability research, and to explore the structure and content of European research programmes relevant for sustainable development. ICT-ENSURE will organise thematic expert workshops to expand the network among experts on environmental sustainability from the EU 27 member states. Selected areas of scientific interest, of economic importance and of political relevance will be analysed by experts to evaluate the importance of ICT in current environmental research, application and education processes and to determine future developments and opportunities.
Development of web-based information systems allowing access to full text paper resources of the environmental sustainability community to content and results generated in the course of the project through expert discussions and scientific surveys and to meta-information concerning European national research programmes in the field of ICT and environment. Based on this ICT-ENSURE will work towards a Single Information Space in Europe for the Environment (SISE). Further information and details can be found on the WebPage: *http://ict-ensure.tugraz.at/content/index.php/ensure/Overview*.

The Environmental Informatics initiative does not have a scientific journal of its own. However, a Special Issue entitled "Environmental Informatics" of the scientific journal Environmental Modelling and Software was published recently by Hilty [2006b]. The history, current research and trends in environmental informatics are given in this issue.

2. 3 Enviromental Modelling and Software

The iEMSs (International Environmental Modelling and Software Society) is a not-for-profit organization uniting private persons and organizations dealing with environmental modelling, software and related topics.
The aims of the iEMSs are to:

- develop and use environmental modelling and software tools to advance the science and improve decision making with respect to resource and environmental issues. This places an emphasis on interdisciplinarity and the development of generic frameworks and methodologies which integrate models and software tools across issues, scales, disciplines and stakeholders with respect to resource and environmental issues;
- promote contacts among physical, social and natural scientists, economists and software developers from different countries and coordinate their activities;
- improve the cooperation between the sciences and decision makers/advisors on environmental matters;
- exchange information in the field of environmental modelling and software among scientific and educational organizations and private enterprises, as well as non-governmental organizations and governmental bodies.

In order to achieve these aims, the iEMSs publishes scientific studies and popular scientific materials in the Environmental Modelling and Software journal (Elsevier); hosts a website which allows members to communicate research and other information relevant to the Society's aims with one another and the broader community; delivers regular information to members through the website and mailing lists. This journal publishes contributions in the form of research articles, reviews, short communications as well as software and data news, on recent advances in environmental modelling and/or software. The aim is to improve our capacity to represent, understand, predict or manage the behaviour of environmental systems at all practical scales, and to communicate those improvements to a wide scientific and professional audience. Only one major topic addressed the term sustainability explicitly. This is the subject integrated assessment and management of systems (river basins, regions etc.) for enhancing **sustainability** outcomes – including linked socioeconomic and biophysical models that may be developed with stakeholders for understanding systems, communication and learning, and improving system outcomes.

The iEMSs hosts biennial international conferences, meetings and courses in environmental modelling and software. The scope of these conferences encompasses **sustainability** research issues which are eminent in the field of environmental modelling and software. During the past conference the iEMSs which took place in Barcelona form July 7th till 10th, 2008 several workshops were dedicated explicitly to sustainability, e.g. Workshop 5, entitled "Qualitative models for sustainability" or Workshop 14 entitled "Research in the field of ICT for Environmental Sustainability". Additionally, an invited talk, given by the author of this paper named "Environmental Protection, Sustainable Development & Risk Management", see Voigt, [2008b]. The next iEMSs conference will most probably take place in Ottawa, July 4th – 6th, 2010.

2. 4 Envirometrics

TIES (The International Environmetrics Society) is a non-profit organization aimed to foster the development and use of statistical and other quantitative methods in the environmental sciences, environmental engineering and environmental monitoring and protection. To this end, the Society promotes the participation of statisticians, mathematicians, scientists and engineers in the solution of environmental problems and emphasizes the need for collaboration and for clear communication between individuals from different disciplines and between researchers and practitioners. The TIES society further promotes these objectives by conducting meetings and producing publications, and by encouraging a broad membership of statisticians, mathematicians, engineers, scientists and others interested in furthering the role of statistical and mathematical techniques in service to the environment. The first TIES Conference took place in Cairo, Egypt in 1989. TIES is an international society and its conferences are being held at several continents of the world. TIES 2008 conference was held in Kelowna, BC, Canada. This year's TIES Conference will be held in Bolgna, Italy, July 5th – 9th, 2009.

Abdel El-Shaarawi the founder of the TIES conferences is optimistic about the future of the field of Environmetrics as quoted in an interview article by Esterby [2003]. Environmetrics has been around for a reasonably long time and has been quite successful expanding in terms of its methods and their range of application. Prof El-Shaarawi is of the opinion that without any doubt, statistics is the broadest field of science in terms of the use of its methods. The same formula for advancement as used by the great statisticians in the past, if used by the current generation, will guarantee a bright future for the field. This formula requires that statistical methods should be developed to address real problems in science and technology, and statisticians should be taught how to effectively communicate with scientific disciplines.
TIES became a section of the International Statistical Institute (ISI) in 2008.
An important background work of TIES is the Encyclopedia of Environmetrics, edited by Abded El-Shaarawi and Walter Piegorsch [2002]. The publication of this major reference work, which includes more than 500 detailed articles, is a landmark for environmetrics. It really defines the subject and indicates its future directions. As can be seen, there is no shortage of real problems for scientists to work on, where their investigations have direct impact on improving the quality of life for the current and future generations.
The official journal of the International Environmetrics Society is named Environmetrics, published by Wiley-Interscience. The Editor-in-chief is Abdel El-Shaarawi [2009]. In this journal selected papers from the TIES conferences are published.

TIES also publishes a newsletter which is available via the Webpage (see Table 1).

2. 5 Partial Order in Environmental Science and Chemistry (Hasse)

Partial order is a discipline of Discrete Mathematics. It has to be regarded with respect to the so-called Hasse diagram technique. The graph which describes a particular poset is called a Hasse diagram. Hasse diagrams got the name from the German mathematician H. Hasse who used them to represent algebraic structures, see Hasse [1967] As Hasse diagrams are the visualization of a mathematical concept, namely of partial order, one has to go back until the end of the nineteenth century, where Dedekind and Vogt Rival [1985] made the first important investigations. Parallel to H. Hasse, the American mathematician G. Birkhoff worked on partial orders and made this mathematical structure popular by his famous book "Lattice theory" [1984]. From the pioneering work of E. Halfon, the concept of partial order was introduced in environmental sciences and chemistry [1983, 1986]. The usefulness of partial order in evaluation of environmental and chemical problems has been applied and extended for more than 20 years in environmental sciences and chemistry by Brüggemann and Carlsen [2006]. Since 1998 regularly workshops about partial order in environmental sciences and chemistry take place (see Table 2). The forthcoming workshop will be held in Ghent, Belgium in 2010.

Table 2. Workshops on Partial Order in Environmental Sciences and Chemistry 1998 - 2008

Year	**Place**	**Title**
1998	Berlin, Germany	Order Theoretical Tools in Environmental Sciences, see Brüggemann [1998]
1999	Roskilde, Denmark	Order Theoretical Tools in Environmental Sciences, see Soerensen [2000]
2000	Berlin, Germany	Order Theoretical Tools in Environmental Science and Decision Systems, see Pudenz [2001]
2001	Iffeldorf, Germany	Order Theoretical Tools in Environmental Sciences, Order Theory Meets Multivariate Statistics, see Voigt [2002]
2002	Roskilde, Denmark	Order Theory in Environmental Sciences, see Soerensen [2004]
2004	Bayreuth, Germany	Partial Orders in Environmental Sciences and Chemistry, see Brüggemann [2005]
2006	Verbania, Italy	Ranking Methods and Multicriteria Decision Analysis in Environmental Sciences, No proceedings
2008	Warsaw, Poland	Multicriteria Ordering and Ranking, see Owsinski [2008]

They gave a first state of art of application of partial order in environmental sciences and chemistry. In the past new tools to apply partial orders in application studies were developed. Furthermore, in view of the increasing community of scientists applying partial orders or deriving new concepts, which are of relevance for applications, there is a need for new software, which is easily applicable, freely available and which can be rather quickly extended.

The first Hasse software was already written in the nineteen eighties and was developed under MS-DOS. In the nineteen ninetieth the software was adapted to the MS-Windows platform. The functionalities were constantly enhanced and improved in the following years. The innovative tool called METEOR (Method of Evaluation by Order Theory) which attempts to resolve the incomparabilities among objects by inclusion of external knowledge is incorporated in the WHASSE program by Brüggemann [2008a], Voigt [2008d]. This software named WHASSE written in Delphi has always been available for scientific purposes free of charge. However, new features (e.g. like the similarity analysis

or the fuzzy partial order etc.) made it necessary to develop a new software written in PYTHON (http://www.python.org/) being more flexible than the older software.
The software PyHasse, which is still under development, is considered as a test version and which is certainly not following the lines of professional software development. The name is selected in analogy to WHasse, ‘Py’ refers to the programming language PYTHON. Besides the central platform, called ‘pyhassemenue2’, there are 17 programs, developed for specific tasks. Here we cannot explain all of the 17 programs. The program was described by the author of the software Rainer Brüggemann at the last Workshop held in Warsaw in November 2008, see Brüggemann [2008b].

2. 6 Chemoinformatics

Chemoinformatics is a fairly new name for a discipline that has been around for quite a while. Different people sometimes give rather different definitions of chemoinformatics. The rather broad and general definition by Johann Gasteiger the editor of the Handbook of Chemoinformatics [2003] is: Chemoinformatics is the application of informatics methods to solve chemical problems. Chemoinformatics is the use of computer and informational techniques, applied to a range of problems in the field of chemistry. These in silico techniques are often used in pharmaceutical companies in the process of drug discovery.
The field of Chemoinformatics was not founded, nor was it formally installed. It slowly evolved from several, often quite humble, beginnings. Scientists in various fields of chemistry struggled to develop computer methods in order to manage the enormous amount of information and to find relationships between the structures and properties of a compound. In the 1960s some early developments became evident which led to a flurry of activities in the 1970s, see Gasteiger [2003]. These days a big issue in this respect is the topic Chemoinformatics and Drug Development.
In Germany conferences on Chemoinformatics have been initiated and held by the CIC (Chemistry-Information-Computer - A division of the German Chemical Society). These conferences have international speakers as well as an international audience http://www.gdch.de/strukturen/fg/cic__e.htm. Last year’s conference 4th Conference on Chemoinformatics together with the 22. CIC-Workshop was held in Goslar, Germany, November 22nd – 24th, 2008.
Important aspects of Chemoinformatics are published in the Handbook as well as in the Textbook of Chemoinformatics Gasteiger [2003]. The field of chemometrics has seen considerable growth in 2006. This growth is evidenced by the fact that significant advances are no longer found in the pages of a few specialty publications but across a wide range of mainstream chemistry and general science journals, even in the Journal of the American Chemical Society JACS, see Agrafiotis, [2007]. So, no specific journal on chemoinformatics exists. However, in a variety of chemistry journals the aspects of chemoinformatics are published, e.g. Journal of Chemical Information and Modelling published by ACS (American Chemical Society).

2. 7 Chemometrics

Chemometrics is the application of mathematical or statistical methods to chemical data. The International Chemometrics Society (ICS) offers the following definition: Chemometrics is the chemical discipline that uses mathematics, statistics, and formal logic:

- To design or select optimum experimental and measurement procedures
- To provide maximum relevant chemical information by analyzing chemical data
- To obtain knowledge about chemical systems.

Chemometrics, however, is a relatively young science that was born at the end of the 1960s and became an independent working field in the 1970s. The name "chemometrics" was first coined by the Swedish scientist Svante Wold in the early 1970s. His cooperation with the American analytical chemist Bruce R. Kowalski, who at the time was working on pattern recognition methods in chemistry, resulted in the foundation of the International Chemometrics Society in 1974. The first German working party on chemometrics was founded in 1984. Around that time chemometric working groups were also established in

many other countries Einax [2004]. The International Chemometrics Society has no active WebPage at the moment. In Germany the working group "Chemometrics and Laboratory Data Processing" within the division of Analytical Chemistry in the German Chemical Society (GDCh) emerged from the former working groups of Chemometrics and Laboratory Automation, which united in 1994. This group focuses on the combination of data acquisition in the laboratory, their statistical evaluation, plausibility checks and the appropriate quality assurance as good analysis practice (*http://www.chemie.uni-jena.de/institute/ac/einax/ak_chemo/ak_chemo*).
Current developments in chemometrics are characterized more by applications than by fundamental work in the statistical and/or mathematical field. In addition to classical problems of analytical chemistry, like the optimization methods, interdisciplinary works are gaining more and more interest. The investigation of the environment, see Einax [1997], of food and pharmaceuticals, see Voigt [2004, 2008c], and chemometric modelling of technological processes and process analytical chemistry should be mentioned. A further important topic is the application of chemometrics to analytical quality assurance.
Concerning environmentally-directed chemometrics the reference work Chemometrics in Environmental Sciences by Einax [1997] should be mentioned. A special issue of the journal Analytical and Bioanalytical Chemistry was issued by Einax [2004] which describes the state of the art as well as current trends in chemometrics.
Important journals in this field are: Chemometrics and Intelligent Laboratory Systems published by Elsevier, Journal of Chemometrics published by Wiley-Interscience, and Journal of Computational Chemistry published by Wiley-Interscience.

2. 8 Green (Sustainable) Chemistry with respect to Computer Science or Mathematics / Green Metrics

Current thinking on sustainable development came out of a United Nations Commission on Environment and Development in 1987 Brundtland Commission [1987], which defined sustainable development as: ‘. . . meeting the needs of the present without compromising the ability of future generations to meet their own needs.’
With respect to chemicals we speak of sustainable chemistry (SC) and/or green chemistry (GC). One important element of sustainable chemistry is commonly defined as chemical research aiming at the optimization of chemical processes and products with respect to energy and material consumption, inherent safety, toxicity, environmental degradability, and so on. An increasing number of assessment systems containing quantitative indicators for these aspects are currently being developed. In addition, however, SC should also address the societal aspect of sustainability. With respect to scientific research, the societal aspect is defined here by two requirements: (1) the assumptions, objectives and implications of chemical research and its technical application should be made more transparent to various societal actors; (2) uncertainty and ignorance should be treated more explicitly in the course of scientific research. Meeting these requirements is necessary in order to lift the division between the allegedly disinterested and non-normative scientific research and the value-laden sphere of societal needs, preferences and decision-making situations. This, in turn, is understood here as a contribution to a more sustainable scientific practice. The societal aspect of SC remains to be recognized more fully in all branches of chemical research. One prerequisite for this is the inclusion of SC into chemical education from the very beginning, see Boeschen [2003].

The US Environmental Protection Agency (EPA) coined the phrase Green Chemistry "To promote innovative chemical technologies that reduce or eliminate the use or generation of hazardous substances in the design, manufacture and use of chemical products" during the early 1990s. Over the last years Green Chemistry has gradually become recognized as both a culture and a methodology for achieving sustainability. The concept of green chemistry was being formulated, by Anastas [1998, 2002] at the US Environmental Protection Agency (EPA), to address the environmental issues of both chemical products and the processes by which they are produced. The guiding principle is the design of environmentally benign products and processes (benign by design) which is embodied in

the 12 Principles of Green Chemistry the essence of which can be reduced to the following working definition. Green chemistry efficiently utilises (preferably renewable) raw materials, eliminates waste and avoids the use of toxic and/or hazardous reagents and solvents in the manufacture and application of chemical products. Concerning the diminishing of waste the so-called e-factor has been generally accepted. This measurement has been introduced more than 10 years ago and can be regarded as a form of green metrics Sheldon [2007].
The design for degradation is an extremely important concept of green chemistry. Chemical products should be designed so that at the end of their function they break down into innocuous degradation products and do not persist in the environment Anastas [1998]. According to Kuemmerer [2007] sustainability should be considered from the very beginning of the design of chemicals. Chemicals are a part of modern life. Products are the main emissions of the chemical and pharmaceutical industries. This makes it difficult to hold them back efficiently. Very often they do not become degraded or fully broken down to water, carbon dioxide and inorganic salts. Often, unknown transformation products are formed in the environment. Therefore, according to the principles of green chemistry, the functionality of a chemical should not only include the properties of a chemical necessary for its application, but also easy and fast degradability after its use. Taking into account the full life cycle of chemicals will lead to a different understanding of the functionality necessary for a chemical. In the present discussion, improvement of synthesis and renewable feedstock are very prominent, whereas the environmental properties of the molecules are somewhat underestimated. In this respect it is of urgent need to take a close look into the theoretical aspects of degradation, e.g. modelling, data-analysis, information systems. The term for these approaches could be green or sustainable modelling. One example of such a green chemicals' modelling approach is a recently published paper by Schenker [2007]. In general global multi-media box models are used to calculate the fate of persistent organic chemicals in a global environment and assess long-range transport or arctic contamination. Currently, such models assume substances to degrade in one single step. In reality, however, intermediate degradation products are formed. If those degradation products have a high persistence, bioaccumulation potential and / or toxicity, they should be included in environmental fate models. The environmental fate model CliMoChem was modified to simultaneously calculate a parent compound and several degradation products.

A variety of activities already exist in the field of green or sustainable chemistry. It goes without saying that only a small selection can be named here. To the knowledge no international society of green or sustainable chemistry exists. In Germany a working group named, sustainable chemistry has been established within the German Chemical Society (GDCh) (*http://www.gdch.de/strukturen/fg/nachhaltigechemie.htm*). The International Symposium on Green Chemistry for Environment and Health was held at the Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, October 13-16, 2008, in Munich, Germany (http://www.helmholtz-muenchen.de/en/gc/).
The mission of the US Green Chemistry Institute, established in 1997, is to advance the implementation of green chemistry principles into all aspects of the chemical enterprise. On January 1, 2001, the ACS Green Chemistry Institute® became part of ACS operations in Washington, D.C. The organization pursues their joint interests in the discovery and design of chemical products and processes that eliminate generation and use of hazardous substances, see ACS [2009]. The Green Chemistry Institute hosts international conferences in this field. This year's event is the 13th Annual Conference on Green Chemistry & Engineering, June 23-25, 2009, College Park, MD (http://acswebcontent.acs.org/gcande/). The following seven different tracks will be offered: Transforming Green Chemistry and Engineering, Lifecycle Analysis and Green Metrics, Toxicology, Environmental Fate and Effects, Renewable Energy Fuels and Feedstocks, Process Design and Optimization, Greenhouse Gas Management and Prevention, Integrative Education in Green Chemistry and Engineering. The Center for Green Chemistry and Green Engineering at Yale headed by Paul Anastas is dedicated to advancing the theory and practice of Green Chemistry and Green Engineering. The Center conducts projects in developing new science, technology, educational opportunities, and policies, for the ultimate goal of increasing the adoption and

implementation of Green Chemistry and Green Engineering throughout our society and our economy for more sustainable world (http://www.greenchemistry.yale.edu/about_the_center/).

3 NETWORKING AMONG the DISCIPLINES

All the discussed disciplines are organized in international groups and societies who work in theoretical fields and aim at making contributions to the better understanding of environmental problems, giving an impact to sustainability and finally aiming at improving the environmental conditions of our planet. All groups are also aware and prepared to do some interdisciplinary work. This interdisciplinarity is – to the knowledge of the author – mainly focussed in the direction of environmental protection and theoretical approaches or in the direction of chemistry, chemicals and theoretical issues and not with respect to the interaction between or among different theoretically-oriented societies and groups.
Only vey few attempts have been made so far to initiate common sessions at conferences or common workshops. At the EnviroInfo 2003, 17th International Conference Informatics for Environmental Protection EnviroInfo in Cottbus, Germany, a workshop on chemometrics was organized by Gnauck [2003]. At the EnviroInfo in Vienna, Austria a TIES Session took place, see Pillmann [2002]. At the EnviroInfo 2007 in Warsaw a track on "Ranking Methods with the emphasis on Partially Ordered Sets" was organized by Brüggemann [2007]. A recent attempt has been made at the iEMSS in Barcelona while organizing a session on Environmental Informatics, History and Current Trends by the author of this paper, see Voigt, [2008a].

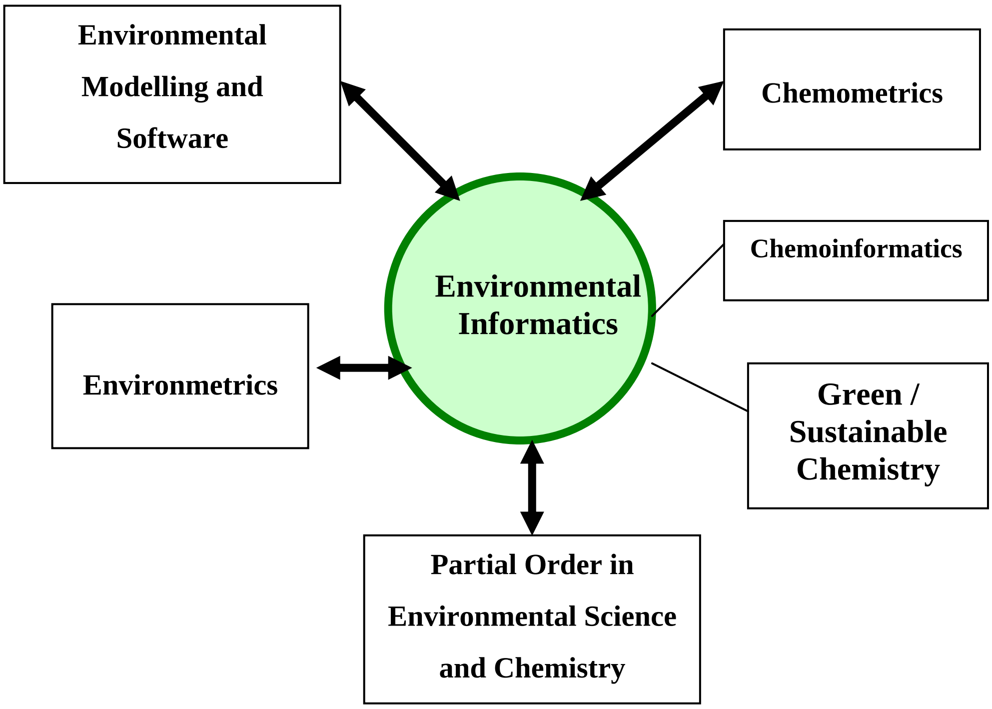


Figure 3. Societies and Initiatives which Focus on Computer Scientific, Mathematical and Statistical Aspects of Environmental Protection Issues

These were single events and did not have the impact on the different societies and groups of experts as expected. So unfortunately these kind of cooperative activities have been discontinued so far. Of course quite a few experts attend several of the mentioned interdisciplinary conferences and present their results.
Figure 3 should demonstrate the current status (known to the author) in the co operations between the named groups and societies.
Currently the societies have more or less one direction: the improvement and protection of our environment by applying their theoretical know-how. For the chemical societies this is only the case in those research areas in which environmental topics, like environmental

chemicals are treated. The aim could be reached more effectively if the cooperation between the groups was intensified.
It is envisaged that the societies and experts continue to improve their efforts in collaboration and taking advantages of other theoretical aspects and ideas concerning environmental protection.

4. FUTURE PERSPECTVES

Humanity has exceeded the carrying capacity of the global environment. The only real choices for the future are to bring the throughputs that support human activities down to sustainable levels through human choice, human technology, and human organization, or to let nature force the decision through lack of food, energy, or materials, or through an increasingly unhealthy environment, see Meadows [2004]. Theoretical approaches describing, modelling and predicting the environmental conditions play an important role in the pursuit of sustainability. One small but important aspect is that the groups, societies which are already working in the field of environmental protection and sustainability increase their knowledge about other activities and group their efforts by improving and intensifying their collaboration in the future. This corresponds to the visioning by Donella Meadows: A sustainable world can never be fully realised until it is widely envisioned. One aspect which is listed is "A flourishing of science, a continuous enlargement of human knowledge", see Meadows [2004. p. 274].

Speaking in practical terms the following aspects should be focussed:

- Taking into consideration the social aspect of sustainability.
- Taking into consideration the economic aspect of sustainability (enviro-economics).
- Including of sustainable issues into the education system.
- Intensifying the information of the public.
- Initiating further research studies in the sustainable modelling, sustainable metrics field.
- Challenging the fine chemical and pharmaceutical industries to make a paradigm shift for the concept of process efficiency which is exclusively focused on the chemical yield to one that is motivated by elimination of waste and maximisation of real material utilisation, see Sheldon [1997].
- Improving collaboration not only among the named societies but also between environmentally and chemically-oriented research and other sciences (e.g. ecomonics).

REFERENCES

ACS (American Chemical Society), ACS Green Chemistry Institute, http://portal.acs.org/portal/acs/corg/content?_nfpb=true&_pageLabel=PP_TRANSITIONMAIN&node_id=1400&use_sec=false&sec_url_var=region1&__uuid=8f601fb5-ca7a-4aca-8626-c52db23e8a01, 2009.

Agrafiotis, D.K., D. Bandyopadhyay, J.K. Wegner, and H. van Vlijmen, Recent Advances in Chemoinformatics, *J. Chem. Inf. Model.*, 47, 1279-1293, 2007.

Anastas, P. T., and T. C. Williamson (Eds), *Green Chemistry: Frontiers in Chemical Synthesis and Processes*, Oxford University Press, Oxford, 1998.

Anastas, P. T., M. M. Kirchhoff, Acc. Chem. Res., 35, 686–693, 2002.

Avouris, N.M., and B. Page, *Environmental Informatics, Methodology and Applications of Environmental Information Processing*, Kluwer Academic Publishers, Dodrecht, 1995.

Birkhoff, G. *Lattice theory*. American Mathematical Society, Vol: XXV, Providence, Rhode Island (USA), 1–418, 1984.

Boeschen, S., D. Lenoir, and Scheringer, M. Sustainable chemistry: starting points and prospects, *Naturwissenschaften*, 90, 93-102, 2003.

Brundtland Commission, *Our Common Future, Report ofn the World Commission on Environment and Development*, http://www.worldinbalance.net/agreements/1987-brundtland.php, 1987.

Brüggemann, R., *Proceedings of the Workshop on Order Theoretical Tools in Environmental Sciences* held on November 16th, 1998 in Berlin, Berichte des IGB 1998, Heft 6, Sonderheft I, Institut für Gewässerökologie und Binnenfischerei, Berlin, 1-117, 1998.

Brüggemann, R., H. Frank., and A. Kerber (Eds.), *Partial Orders in Environmental Sciences and Chemistry*, match, Communications in Mathematical and in Computer Chemistry, 54, 3, 1487-689, 2005.

Brüggemann, R., and L. Carlsen, *Partial Order in Environmental Sciences and Chemistry*, Springer-Verlag, Berlin (GE), 1–406, 2006.

Brüggemann R., K. Voigt, P.B. Sørensen, and B. De Baets, Partial Order Concepts in Ranking Environmental Chemicals, in: Hryniewicz Olgierd, Studzinski Jan, Szediw Anna, *EnviroInfo Warsaw 2007, Environmental Informatics and Systems Research*, Volume 2: Workshop and application papers, pp. 169-176, Shaker Verlag, Aachen, 2007.

Brüggemann R., Voigt K., Restrepo G., and Simon U., Concept of Stability Fields and Hot Spots in Ranking of Environmental Chemicals, *Environmental Modelling & Software*, 23, 1000-1012, 2008a.

Brüggemann, R., G. Restrepo, and K. Voigt., Towards a new and advanced Partial Order Program: PYHASSE in: Owsinski, J., Brüggemann R. (Eds.), *Multicriteria Ordering and Ranking: Partial Orders, Ambiguities and Applied Issues*, 11-34, Systems Research Institute, Polish Academy of Sciences, Warsaw, 2008b.

Einax, J., H.W. Zwanzinger, and S. Geiss, *Chemometrics in Environmental Analysis*, VCH Verlagsgesellschaft, Weinheim, 1997.

Einax, J., *Chemometrics in Analytical Chemistry, Analytical & Bioanalytical Chemistry*, 380 (3), 368-369, 2004.

El-Shaarawi, A., and W. Pigorsch, *Encyclopedia of Environmetrics*, Wiley, Weinheim, 2002.

El-Shaarawi, A., (Editor in Chief), *Environmetrics*, the official journal of the International Environmetrics Society (TIES), Wiley-Interscience, 2009.

Esterby, S., A Conversation with Abdel H. El-Shaarawi, *Statistical Science*, 18 (4), 481-488, 2003.

Gasteiger, J., and T. Engel, *Chemoinformatics, A Textbook and Handbook*, Wiley-VCH, Weinheim, p. 4, 2003.

Gnauck, A., and R. Heinrich, *EnviroInfo 2003, The Information Society and the Enlargement of the European Union*, 17th International Conference Informatics for Environmental Protection, Part 2: Applications, Workshops, Posters, Metropolis Verlag, Marburg, 2003.

Halfon, E. Is there a best model structure? I: Modelling the fate of a toxic substance in a lake, *Ecol. Model.*, 20, 135–152, 1983.

Halfon, E., and Reggiani, M.G, On ranking chemicals for environmental hazard, *Environ. Sci. Technol* , 20, 1173–1179, 1986.

Hasse, H., *Vorlesungen ueber Klassenkoerpertheorie*, Physica-Verlag, Germany, 275, 1967.

Hepting, D.H. Decision support for local environmental impact assessment, *Environmental Modelling & Software*, 22, 436-441, 2007.

Hilty, L.M., *Umweltbezogene Informationsverarbeitung*, Beiträge der Informatik zu einer nachhaltigen Entwicklung. Habilitationsschrift am Fachbereich Informatik der Universität Hamburg, 1997.

Hilty, L.M., and Gilgen, P.W. (Eds), *Sustainability in the information society*. 15th International Symposium Informatics for Environmental Protection. Zurich 2001. Part 1 and Part 2. Marburg: Metropolis, 2001.

Hilty, L.M., E. Seifert., and R. Treibert. (Eds.), I*nformation Systems for Sustainable Development*., Idea Group Publishing, Hershey (PA), 2004.
Hilty, L.M., P. Arnfalk, L. Erdmann, J. Goodman, M. Lehmann, and P. Waeger., The relevance of information and communication technologies for environmental sustainability: a prospective simulation study, *Environmental Modelling & Software*, 21 (11), 1618-1629, 2006a.
Hilty, L.M., B. Page, and J. Hrebicek (Guest Editors), Special Issue *Environmental Informatics, Environmental Modelling & Software*, 21 (11) 1517-1629, 2006b.
Isenmann, R., Environmental Informatics and Industrial Ecology – Scientific Profiles of Two Emerging Fields Striving for Sustainability, In: Moeller, A., B. Page, M. Schreiber, E*nviroInfo 2008, Environmental Informatics and Industrial Ecology* 22nd International Conference on Informatics for Environmental Protection, Shaker Verlag, Aachen, 14-22, 2008.
Kuemmerer, K., Sustainable from the very beginning: rational design of molecules by life cycle engineering as an important approach for green pharmacy and green chemistry, *Green Chemistry*, 9, 899-907, 2007.
Meadows, D., J. Randers, and D. Meadows, *Limits to Growth, the 30-year update*, Chelsea Green Publishing Company, White River Junction, Vermont, 2004.
OECD, *Towards Sustainable Household Consumption? Trends and Policies in the OECD Countries*, OECD, Paris, 2002.
Owsinski, J. W., and Brüggemann R. (Eds.), *Multicriteria Ordering and Ranking: Partial Orders, Ambiguities and Applied Issues*, Systems Research Institute, Polish Academy of Sciences, Warsaw, 1-269. 2008.
Page, B., A. Jaeschke, and W. Pillmann, Angewandte Informatik im Umweltschutz. Teil 2. *Informatik Spektrum*, 13: 86-97, 1990.
Page, B., and L.M. Hilty, *Umweltinformatik, Informatikmethoden fuer Umweltschutz und Umweltforschung*, 2. aktualisierte und erweiterte Auflage, R. Oldenbourg Verlag, Muenchen, 1995.
Pillmann, W., and K. Tochtermann, *EnviroInfo 2002, 16th International Conference, Informatics for Environmental Protection, Environmental Communication in the Information Society*, International Society for Environmental Protection, Vienna, 2002.
Pillmann, W., W. Geiger, and K. Voigt, Survey of Environmental Informatics in Europe, *Environmental Modelling & Software*, 21 (11), 1519-1527, 2006a.
Pillmann, W., W. Geiger, and R. Isenmann, Informatics for Environmental Protection, Sustainable Development, and Risk Management, 20 Years Technical Committee on environmental Informatics In: Tochtermann, K. and A Scharl (Eds.), *EnviroInfo 2006, 20th International Conference on Informatics for Environmental Protection, Managing Environmental Knowledge*, 1-14, 2006b.
Pudenz, S., R. Brüggemann, and H.-P. Lühr (Eds), *Order Theoretical tools on Environmental Science and Decision Systems*. Proceedings of the third Workshop November 6th – 7th, 2000 in Berlin Germany, Leibniz Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany, - 222 – Heft 14, Sonderheft IV, 1-224, 2001.
Rautenstrauch, C., and S. Patig, *Environmental Information Systems in Industry and Public Administration*, Idea Group Publishing, London, 2001
Rival, I., *The diagram. In: Graphs and Order* (Rival, I. ed.), D. Reidel Publishing Company, Dordrecht (NL), 103–133, 1985.
Schenker, U., M. Scheringer, and K. Hungerbuehler, Including Degradation Products of Persistent Organic Pollutants in a Global Multi-Media Box Model., *Env Sci Pollut Res*. 14, 145 – 152, 2007.
Sheldon, R. A., The E Factor: fifteen years on, *Green Chemistry*, 9, 1273-1283, 2007.
Sørensen, P.B., L. Carlsen, B.B. Mogensen, R. Brüggemann., B. Luther, S. Pudenz., U. Simon, E. Halfon, K. Voigt, G. Welzl, and G. Rediske, *Order Theoretical Tools in Environmental Sciences*, Proceedings of the Second Workshop October 21st, 1999 held in Roskilde, Denmark, NERI – Technical Report No. 318,. National Environmental Research Institute, Roskilde, Denmark, 1-170, 2000.
Sørensen, P., Brüggemann R., Lerche D.B., and Voigt K., *Order Theory in Environmental Sciences*, NERI Technical Report No. 479, , pp. 52-67, National Environmental Research Institute, Ministry of the Environment, Copenhagen, 2004.

Voigt K., and G. Welzl. (Eds.), *Order Theoretical Tools in Environmental Sciences, Order Theory (Hasse Diagram Technique) Meets Multivatiate Statistics*, Shaker-Verlag, Aachen, 1-206, 2002.

Voigt K., R. Brüggemann R., and S. Pudenz, Chemical Databases Evaluated by Order Theoretical Tools, *Analytical and Bioanalytical Chemistry*, 380, 467-474, 2004.

Voigt K., Environmental Informatics, Environmetrics, Chemoinformatics, Chemometrics: Integration or Separation!? In: Sànchez-Marrè M. et al. (Eds.), *Proceedings of iEMSs 2008*, Barcelona, Catalonia, July 2008, Volume 3, pp. 1594-1601, 2008a.

Voigt K., Environmental Protection, Sustainable Development & Risk Management with Emphasis in Environmental Chemicals. in: Sànchez-Marrè M. et al. (Eds.), *Proceedings of the iEMSs 2008*, Barcelona, Catalonia, July 2008, Volume 1, pp. 7-11, 2008b.

Voigt, K., and R. Brüggemann., Ranking of Pharmaceuticals Detected in the Environment: Aggregation and Weighting Procedures, C*ombinatorial Chemistry & High Throughput Screening*, 11, 10, 770-782, 2008c.

Voigt, K., R. Brüggemann., K.-W. Schramm, and M. Kirchner, PYHASSE a New Software Tool for Partially Ordered Sets: Ranking Soil against Needle Samples, in: Jan W. Owsinski, Rainer Brüggemann (Eds.), *Multicriteria Ordering and Ranking: Partial Orders, Ambiguities and Applied Issue*s, 35-45, Systems Research Institute, Polish Academy of Sciences, Warsaw, 2008d.

Workshop 3

eEcoRA - Information Technologies in Ecological Risk Assessment

Organized by **Ivan Holoubek**, Karel Bláha,
Ladislav Dušek and Alberto Susin

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

An Intelligent System for Monitoring and Predicting Water Quality

N. Bassiliades[a], I. P. Antoniades[b], E. Hatzikos[c], I. Vlahavas[a,b], C. Koutitas[b,d]

[a] *Department of Informatics, Aristotle University of Thessaloniki, Greece (nbassili@csd.auth.gr, vlahavas@csd.auth.gr)*

[b] *International Hellenic University, Thessaloniki, Greece*

[c] *Department of Automation, Technological Education Institution (TEI) of Thessaloniki, Greece (hatzikos@teithe.gr)*

[d] *Department of Civil Engineering, Aristotle University of Thessaloniki, Greece (koutitas@civil.auth.gr)*

Abstract: In this paper we present an intelligent system for monitoring and predicting water quality, whose main aim is to help the authorities in the "decision-making" process in the battle against the pollution of the aquatic environment, which is very vital for the public health and the economy of Northern Greece. Two sensor-telematic networks for collecting water quality measurements in real time (Andromeda, for sea waters, and Interrisk, for surface/fresh waters) were developed and deployed. Sensor readings (water temperature, pH, dissolved oxygen, conductance, turbidity, sea currents, and salinity) are transmitted to a main station for processing and storage. The intelligent system monitors sensor data, reasons, using fuzzy logic, about the current level of water suitability for various aquatic uses, such as swimming and piscicultures, and flags out appropriate alerts. Furthermore, the system employs Machine Learning and Adaptive Filtering techniques and algorithms which successfully predict measurements a day ahead, as well as techniques to incorporate the window of past values in order to be able to make a more precise prediction. The results showed that these algorithms can help make accurate predictions one day ahead and are better than the naive prediction that the value will be similar to today.

Keywords: Sensor Network; Pollution Monitoring; Aquatic Uses; Water quality; Pollution Prediction; Knowledge-Based System.

1. INTRODUCTION

The environment includes the atmosphere, the soil and the water. By the term "water", we mean in general the aquatic resources, either surface waters (e.g. seas, lakes, tanks, rivers) or underground aquatic volumes (Bing et al. [2002]). In the technologically advanced countries, particularly in the USA and the countries of Western Europe, the need for monitoring the parameters related with qualitative environmental characteristics and especially with the quality of water, has been recognized for a long time. To this aim, several programs of automatic measurement of qualitative characteristics and analysis of results have been installed and placed in operation. If the authorities or organizations involved in the management of aquatic resources were able to monitor the quantitative and qualitative parameters related to the aquatic environment, they would be able to draw conclusions about trends, to predict undesirable situations on time and, therefore, to take counter-measures in a timely fashion. Furthermore, they could enforce longer term actions, including strategic resource planning for regional growth and development.

Research groups around the world are dealing with issues concerning environmental pollution. Considering that prevention is the best way to fight pollution, this paper deals with an intelligent system for the permanent monitoring of the quality of water as well as atmospheric and meteorological data. Up to today the main goal of research was the collection of data in real time. It soon became obvious that there was a lack of dynamic systems that would have the capability of processing the data in conjunction with knowledge bases capable to support decision making systems, risk analysis and early warning. At the same time these systems should include prediction capabilities to predict the evolution of phenomena's as well as defining future steps.

In the scope of this work two sensor-telematic networks for collecting water quality measurements in real time (Andromeda, for sea waters, and Interisk, for surface/fresh waters) were developed and deployed. Sensor readings (water temperature, pH, dissolved oxygen, conductance, turbidity, sea currents, and salinity) are transmitted to a main station for processing and storage.

Having numerous data collected on-line, we were able to develop an intelligent system that not only monitors the environmental pollution through assessing the water quality for various aquatic uses, but also includes early warning techniques to predict upcoming natural disasters. This is achieved via an expert system that monitors water quality and pollution (Hatzikos et al. [2007]). The expert system monitors sensor data collected by Local Monitoring Stations (LMS) and reasons about the current level of water suitability for various aquatic uses, such as swimming and piscicultures. The aim of the expert system is to help the authorities in the "decision-making" process in the battle against the pollution of the aquatic environment, which is very vital for the public health and the economy of Northern Greece. The expert system determines, using fuzzy logic, when certain environmental parameters exceed certain "pollution" limits, which are specified either by the authorities or by environmental scientists, and flags out appropriate alerts.

Furthermore, the system employs Machine Learning (Hatzikos et al. [2008]) and Adaptive Filtering (Hatzikos et al. [2009]) techniques and algorithms which successfully predict measurements a day ahead, as well as techniques to incorporate the window of past values in order to be able to make a more precise prediction. More specifically, the data acquired by the system have been used to study the problem of water quality prediction, performing both exploratory and automatic analysis of the collected data with a variety of methods. The results showed that these algorithms can help make accurate predictions one day ahead and are better than the naive prediction that the value will be similar to today.

The rest of the paper is structured as follows: Section 2 briefly reviews relevant systems found in the literature. Section 3 presents the overall system architecture, as well as issues related to the hardware and the software involved in building the telematic sensor networks. Section 4 presents the expert system that monitors water pollution levels using fuzzy logic. Section 5 briefly overviews results obtained for predicting water quality measurements using various techniques, presenting adaptive filtering techniques in more detail. Finally, Section 6 concludes this paper, summarizing its major issues and presenting ground for future work.

2. RELATED WORK

There exist quite a few systems in the literature that monitor the environment, in general, or sea/fresh water, in specific, and alert the user about possible dangers or water suitability. Usually, such systems are strongly related to the local environment they monitor, since environmental monitoring is a very complex task that is strongly dependent on the geomorphologic features of the monitored area. To the best of our knowledge, our system is the first one for monitoring the aquatic environment of northern Greece. In this section we review few such environmental and/or water monitoring systems.

In Hendee [1998], an expert system for marine environmental monitoring is presented, through the SEAKEYS network, which is situated along 220 miles of the coral reef tract within the Florida Keys National Marine Sanctuary. This network monitors meteorological parameters (wind speed, wind gusts, air temperature, barometric pressure, relative humidity), along with oceanographic parameters (sea temperature, photosynthetically active radiation, salinity, fluorometry, optical density). Then, an expert system is employed to provide daily interpretations of near real-time acquired data for the benefit of scientists, fishermen and skin divers. These interpretations are designed to be automatically emailed to Sanctuary managers of the network. The initial set of interpretations included environmental conditions conducive to coral bleaching.

Another knowledge-based approach was in Saunders et al. [2005], to build a system whose main goal is to classify lake-water resources in five acid-sensitive regions of the United States. It consists of a network of Decision Support Systems (DSS), one for each region. It is based on a set of rules, which is knowledge acquired from human acid-base chemistry experts. The DSS allows federal land managers to conduct a preliminary assessment of the status of individual lakes prior to consulting an expert. The authors claim that the DSS accurately portrays the decision structure and assessment outcomes of domain experts while capturing interregional differences in acidification sensitivity and historic acid deposition loadings. It is robust with respect to missing water chemistry input data.

The case-based reasoning system, presented in Fdez-Riverola and Corchado [2003], specializes in forecasting the red tide phenomenon in a complex and dynamic environment in an unsupervised way. Red tides are sea water discolorations caused by dense concentrations of phytoplankton. The system is an autonomous Case-Based Reasoning (CBR) hybrid system that embeds various artificial intelligence tools, such as case-based reasoning, neural networks and fuzzy logic in order to achieve real time forecasting. It predicts the occurrence of red tides caused by the pseudo-nitzschia spp diatom dinoflagellate near the North West coast of the Iberian Peninsula. Its goal is to predict the pseudo-nitzschia spp concentration (cells/liter) one week in advance, based on the recorded measurements over the past two weeks. The Authors claim that their prototype forecasts with an acceptable degree of accuracy. The results obtained may be extrapolated to provide forecasts further ahead. However, the further ahead the forecast is made, the less accurate it may be.

In Cheng et al. [2003] an expert system is presented for assisting the improvement of water quality in a city, applied in the Yellow River Basin of China. The system can analyze relationships between industrial water pollution and economic activities of industrial enterprises of a city. The system includes a decision model at its core, which integrates another four closely related subsystems. According to the authors, the system could provide better decision support for environmental management.

Finally, in Lee et al. [1997] a fuzzy expert system is presented for the determination of Water Quality Classification for Stream (WQCS) from uncertain and imprecise ecological information. The system employs 30 rules, generated from a rule matrix of seven water quality grades, toxicity of water and rarity of cases. According to paper results, smoothly varying curves of WQCS determination from the fuzzy expert system represented real-world experience more realistically than stepwise curves from a conventional expert system.

The work related to water quality prediction includes a variety of linear and nonlinear modelling techniques. Among the various models used are the early Bayesian probability network models focusing both on their accuracy and the correct characterization of the processes (Reckhow [1999]), the predictive clustering approach using a single decision tree for simultaneous prediction of multiple physico-chemical properties of river water from its current biological properties (Blockeed et al. [1999]), the work of regression trees for predicting chemical parameters of river water quality from bioindicator data (Dzeroski et al. [2000]), and the unvaried time series models for determining the long-term and seasonal behaviour of important water quality parameters (Lehmann and Rode [2001]).

3. SYSTEM ARCHITECTURE

The general architecture of both the Andromeda and the Interrisk sensor networks is shown in Figure 1. Each network consists of LMSs that host sensor plunged into water and collect aquatic numeric data concerning sea or fresh water. Sensor readings are transmitted to a Main Station (MS) for processing and storage.

In the case of the Andromeda sea water network, the 3 LMSs are plunged into Thermaikos Gulf and the following hydrological parameters are measured: water temperature, pH, dissolved oxygen, conductance, turbidity, sea currents, and salinity. The transmission of data between the LMSs and the MS is done via radio modems. A more thorough description can be found in Hatzikos [2002].

The Interrisk network collects, from lakes Doirani and Kerkini and the Strimonas River, hydrological parameters for fresh water (water temperature, pH, dissolved oxygen, % oxygen, turbidity, conductivity, water depth), as well as, meteorological parameters (air temperature, air relative humidity, solar radiation, wind speed, wind direction, rainfall, evaporation). The communication of the LMSs with the MS is performed through mobile telephony (GSM technology) with the use of suitable GSM Modems.

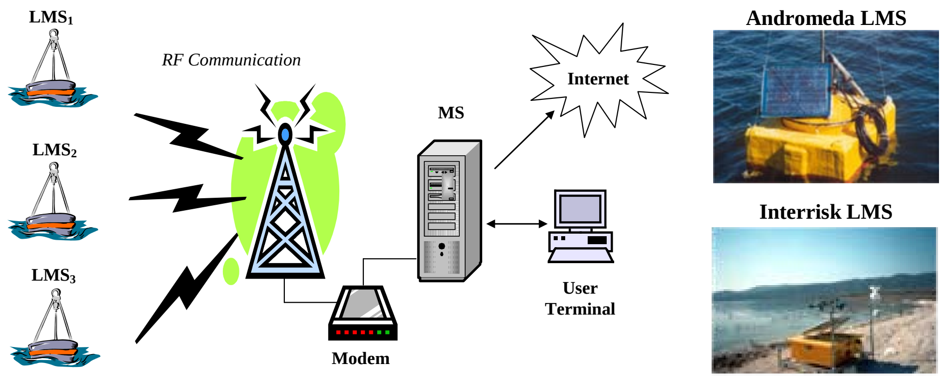


Figure 1. Architecture of the Sensor Networks.

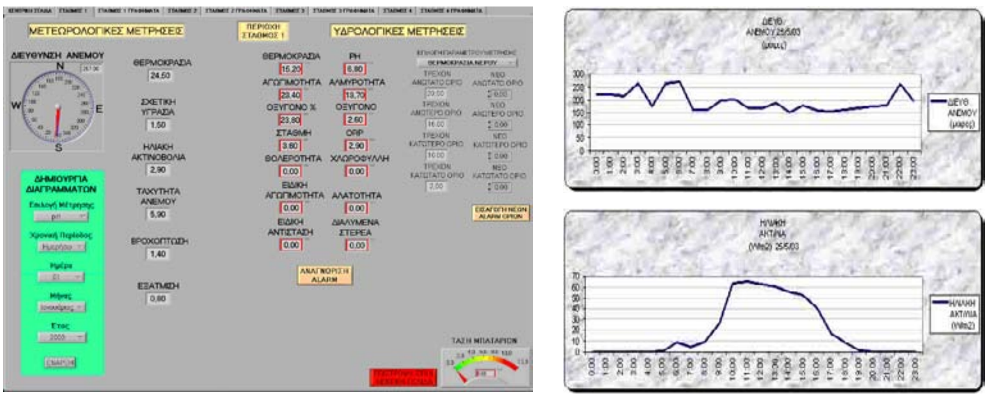


Figure 2. Sensor data visualization at the main station.

The Main Station (MS) is a workstation that collects sensor measurements from all the LMSs and visualizes the results in a SCADA environment. The MS initiates the communication with each of the LMSs in predetermined time intervals using a hand-shake technique. The MS also adjusts the frequency of measurements depending on the situation at hand, i.e. an emergency in the case of pollution. The LMS operates only during the rendezvous. In this way, less energy is consumed. Furthermore, the on-demand measurement policy achieves a higher level of flexibility.

The data collection from the sensors and its transmission to the MS is performed via the SCADA software. LabView is responsible for the data reception, visualization (Figure 2), long-term storage in databases and exporting them in various formats for further processing. Additionally, the administrator of the MS can set alarms in LabView (to be received by the user) when certain sensor measurements exceed some predefined limits.

4. MONITORING WATER QUALITY THROUGH A FUZZY EXPERT SYSTEM

As mentioned in the previous section, the check whether a sensor measurement exceeds certain pollution limits is performed in a rigid fashion, not allowing flexibility. To this end, the MS is equipped with a fuzzy expert system, implemented in MATLAB, which alleviates this by using fuzzy logic to check for violation of environmental parameter limits. The expert system is fed with data from the SCADA software and produces results about the water suitability regarding swimming, shell-culture and pisciculture, also alerting the user. Expert scientific knowledge is required for these issues to be resolved. If a human expert is not available at the corresponding operation centre, then the hydrological parameter values should be monitored automatically by a corresponding expert system, in order to draw conclusions about possible hazardous situations for the environment.

The scientific knowledge required for the expert system was elicited from the Greek environmental legislation for the Region of Central Macedonia. The desired and allowed values for the various measurements are set by the aforementioned law depending on the aquatic use: drinking water, swimming, shell cultivation, etc. This particular expert system deals with shell cultivation, swimming, Cyprinidae cultivation and Salmonidae cultivation.

All the above ranges are not crispy but fuzzily defined, in order to allow for fluctuations in sensor readings, due to either limited sensor accuracy or random fluctuations of physical conditions. Furthermore, fuzzy ranges can meet ascending or descending parameter trends, in the absence of proper trend analysis within the system. For example, if there is a descending trend for the pH that would turn it from normal to acidulous in a few hours, fuzzy ranges would cause an alert before the pH actually reaches the critical point. This means that authorities will be alerted in time to be prepared for action, when the actual critical situation arises. Notice that this approach might also cause false alarms due to random sensor reading fluctuations, but the alerting and mitigations strategies, i.e. importance and frequency of the alerts and proper measures to be taken, is not a part of the monitoring system.

A final justification for using fuzzy logic is that the latter can cater for the rigidness imposed by the legislation, which had to be formulated precisely, using crispy values. However, in practice such limits are never rigidly defined, since they are derived statistically. Notice that the overlapping of the fuzzy ranges are narrow, since the legislation cannot be highly disputed, meaning that measures against pollution and/or polluters cannot be taken if the actual crispy limits set by the law are not violated. Having large overlaps would cause more pollution alarms that would trigger authorities unnecessarily. In any case, the degree of overlapping was actually established through experimentation and it was tuned for smooth behaviour of the Fuzzy Inference System (FIS) transition function.

There is one FIS for each hydrological parameter. Each FIS has one input variable and four output variables, one per aquatic use. The use of multiple single-input-multiple-output (SIMO) FISs instead of one multiple-input-multiple-output (MIMO) FIS reduces the complexity of the transition function and is justified by the fact that the legislation correlates non-acceptable parameter values with a single water usage, i.e. it regards each variable as independent from the other ones.

The input variable is split in as many membership functions as needed to utilize the desired and allowed ranges for each aquatic use (Figure 3). The output variables affected by this particular measurement contain three membership functions: OK, CAUTION and DANGER. In each FIS, if the input falls within the desired range for a particular use, the respective output for this use is OK. In case the input is outside the desired range of values, but remains within the allowed range for a particular use, the corresponding output is CAUTION, otherwise it is DANGER. The rules for each FIS (Figure 4) were formulated by taking into account: a) the correlation between the input and the output variables, and b) the definitions of the membership functions. The system uses the Mamdani's min fuzzy inference method and the centroid defuzzification method.

Finally, overall conclusions are presented to the user. The system must be able to provide answers to questions like the following: "*Is the water suitable for swimming*?" Thus, the user reads on the screen a combination of the aquatic use, the alert type, and the precise measurements that were deemed abnormal (Figure 5).

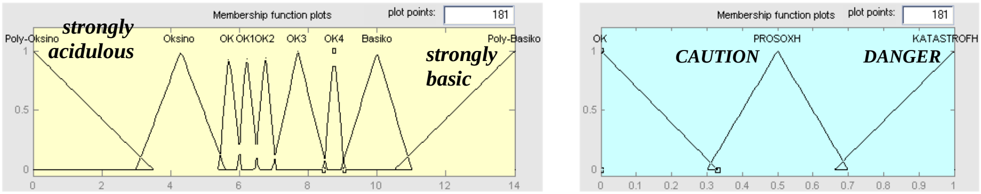


Figure 3. Membership functions of the input variable "pH" and the output variable "Shell".

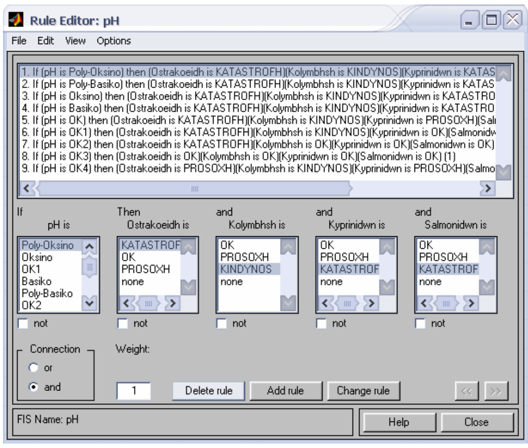
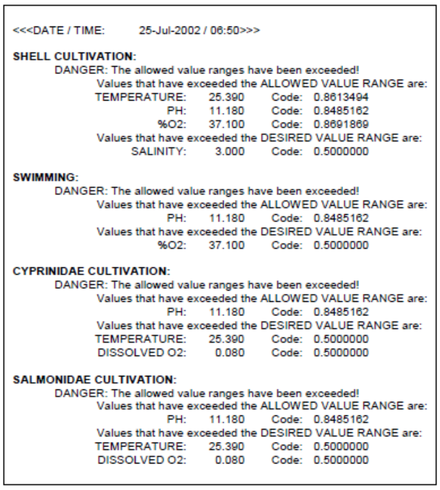


Figure 4. The set of rules of the pH FIS.



<<<DATE / TIME: 25-Jul-2002 / 06:50>>>

SHELL CULTIVATION:
DANGER: The allowed value ranges have been exceeded!
Values that have exceeded the ALLOWED VALUE RANGE are:
TEMPERATURE: 25.390 Code: 0.8613494
PH: 11.180 Code: 0.8485162
%O2: 37.100 Code: 0.8691869
Values that have exceeded the DESIRED VALUE RANGE are:
SALINITY: 3.000 Code: 0.5000000

SWIMMING:
DANGER: The allowed value ranges have been exceeded!
Values that have exceeded the ALLOWED VALUE RANGE are:
PH: 11.180 Code: 0.8485162
Values that have exceeded the DESIRED VALUE RANGE are:
%O2: 37.100 Code: 0.5000000

CYPRINIDAE CULTIVATION:
DANGER: The allowed value ranges have been exceeded!
Values that have exceeded the ALLOWED VALUE RANGE are:
PH: 11.180 Code: 0.8485162
Values that have exceeded the DESIRED VALUE RANGE are:
TEMPERATURE: 25.390 Code: 0.5000000
DISSOLVED O2: 0.080 Code: 0.5000000

SALMONIDAE CULTIVATION:
DANGER: The allowed value ranges have been exceeded!
Values that have exceeded the ALLOWED VALUE RANGE are:
PH: 11.180 Code: 0.8485162
Values that have exceeded the DESIRED VALUE RANGE are:
TEMPERATURE: 25.390 Code: 0.5000000
DISSOLVED O2: 0.080 Code: 0.5000000

Figure 5. Sample expert system output.

5. PREDICTING WATER QUALITY THROUGH ADAPTIVE FILTERING AND MACHINE LEARNING

Over the last decade monitoring water quality and other environmental variables has become considerably important, in an effort to predict their future behavior and prevent undesirable environmental situations, as well as, to enforce longer term actions for regional growth and development. The ability to predict the quality of water in an ecosystem one or more days ahead is very useful, giving the possibility to the authorities for the necessary precautionary actions in time.

In the MS of our sensor networks, we employ both Machine Learning and Adaptive Filtering techniques and algorithms in order to predict measurements a day ahead. In the case of Machine Learning, we also incorporate the window of past values in order to make a more precise prediction. The results showed (Hatzikos et al. [2008]) that the Machine Learning algorithms are able to predict more accurately several days ahead and the furthest ahead the prediction, the largest the window of past values should be incorporated in the model. However, in this paper we focus more on the results of the Adaptive Filtering techniques (Hatzikos et al. [2009]) which are more intuitive.

More specifically, we investigate the possibility to predict a number of water quality variables that are obtained by an under-water measurement set-up. Our interest is focused on one-day ahead predictions of certain water quality variables recorded by an under-water set of sensors, such as water temperature, pH, conductivity, salinity, amount of dissolved oxygen and turbidity. The measured data, forming time series, are stored in a database and a number of modeling methods could then be used to reveal any hidden information. Here we present results obtained for the water temperature, pH, amount of dissolved oxygen and turbidity, due to their higher importance in terms of commercial exploitation.

Adaptive Filtering is performed via a projection and a least squares algorithm. Their prediction ability is shown by comparing their performance against the delayed or naive prediction algorithm (Random Walk Model), which serves as a benchmark model in prediction tasks. The delayed prediction model states simply that the future value (tomorrow) of a variable will be equal to its current value (today), supporting in that way the unpredictability of the modeling object. However, due to the correlation and interaction between the water quality variables, it is interesting to investigate if there is an underlying mechanism that governs the data and thus will prove the predictability of these variables. The projection and least squares algorithms were chosen because they are widely used prediction algorithms and by applying both we can compare and contrast results The identification of such models is particularly useful for ecologists and environmentalists since they will be able to predict in advance the pollution levels in the sea water and thus to instruct all the necessary precaution actions.

The original data had a sampling time of 9 seconds. However, due to the fact that there were a large number of outliers in the data, and that the data did not vary too much on an hourly basis, it was decided that the data should be averaged over a period of one day. Figure 6 shows the autocorrelation sequences for temperature, pH, oxygen and turbidity, calculated from one-month data sets collected during July 2004. This figure demonstrates that temperature, oxygen and turbidity are correlated with past values, and therefore it is possible to use them in prediction. The autocorrelation sequence for pH, on the other hand, is very sharply peaked around time lag 0, which implies that it could be very difficult to construct one-day ahead predictions for this variable.

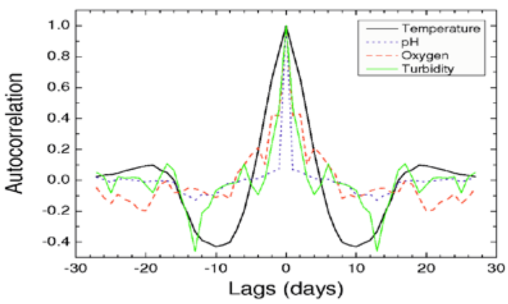


Figure 6. Autocorrelation sequences for temperature, pH, oxygen and turbidity.

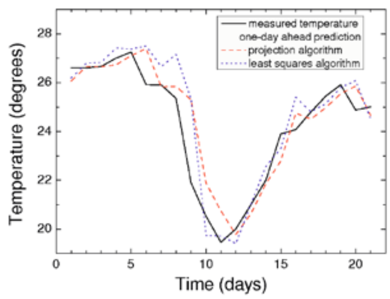


Figure 7. Temperature prediction.

In the following, we show the results we obtained for water temperature prediction, which are more encouraging. Results for the rest of the variables are obtained similarly and can be found in (Hatzikos et al. [2009]).

Initially, in the water temperature prediction, only the values of previous days were used. This was due to the fact that the water temperature is mainly influenced by external variables such as the amount of radiation from the sun, wind direction, etc. Based on the autocorrelation analysis, it was decided that the values of water temperature of the current day and two previous days should be used to predict the water temperature of the next day. The initial estimate was that in most cases the temperature tomorrow should be more or less equal to the temperature today. Before building the prediction model, the data was normalized to zero mean and unity variance. Figure 7 shows the measured and the predicted temperature by use of the Projection algorithm and of the Least squares algorithm.

The prediction of the projection algorithm is quite accurate, in particular after day 12. The algorithm converges in a way that shows that its prediction is based mostly on the current temperature, but corrects the value using previous temperatures as well. The prediction accuracy of the least squares algorithm is reasonable, although not as accurate as the projection algorithm. The algorithm converges in a way that shows that the prediction model uses the current temperature more heavily than the projection algorithm. The observation that the projection algorithm predicts more accurately after day 12 is due to the fact that the measured temperature shows a smoother trend after that day.

Table 1 shows the l_2-norm and l_∞-norm of the prediction error for the projection algorithm, the least squares algorithm and the “delayed prediction algorithm”. Prediction models from both the projection algorithm and least squares algorithm give better prediction accuracy, which is a remarkable result. Furthermore, the prediction model from the projection algorithm is the most accurate one.

Table 1. Prediction accuracy of temperature

Prediction method	l_2	l_∞
Projection algorithm	2.0	1.47
Least square algorithm	2.0	1.5
Delayed prediction algorithm	2.2	5.5

6. CONCLUSIONS AND FUTURE WORK

In this paper we have presented an intelligent system for monitoring and predicting water quality, whose main aim is to help the authorities in the "decision-making" process in the battle against the pollution of the aquatic environment, which is very vital for the public health and the economy of Northern Greece, and not only. The system is realized via two sensor-telematic networks for collecting water quality measurements in real time (Andromeda, for sea, and Interrisk, for fresh waters). Sensor readings are transmitted to a main station for processing and storage.

The intelligent system monitors sensor data, reasons, using fuzzy logic, about the current level of water suitability for various aquatic uses, such as swimming and piscicultures, and flags out appropriate alerts. Furthermore, the system employs Machine Learning and Adaptive Filtering techniques and algorithms which successfully predict measurements a day ahead, as well as techniques to incorporate the window of past values in order to be able to make a more precise prediction. The results showed that these algorithms can help make accurate predictions one day ahead and are better than the Naive prediction that the value will be similar to today.

The main advantage of the system and its architecture is its versatility by means of extensibility and mobility. Concerning the sensor network, new sensors for a variety of environmental readings (e.g. hydrological, meteorological, etc.) can be and have been easily added to the system. Furthermore, existing LMSs can be easily moved to different locations and new LMSs can be easily added, without disturbing the rest of the system. The communication between the LMSs and the MS can be and has been implemented with a variety of technologies, depending on the geomorphologic and socioeconomic features of the installation area. Another advantage of the flexibility of the system is that new methodologies and techniques both for predicting and monitoring can be used without disturbing the rest of the system.

The Andromeda network was working productively from 1998 until 2005, when it ceased working due to lack of Governmental funding. The monitoring system has been working 18 months during 2004-2005. During that period a large number of “pollution events" were recorded, for which the system responded issuing alerts properly, since the LMSs were installed near the port and the industrial area of Thessaloniki, where sea quality is very poor. However, only one severe event was recorded when all sensors indicated strongly a very large deterioration of water quality, which turned out to be due to a spill from a ship. Finally, there were also some false and/or missed alarms occurring on days when the maintenance of sensors was inadequate, or when litter was cluttering sensor readings, in both cases resulting in untrustworthy measurements.

One of the first future research priorities is to integrate the prediction algorithms employed in the MS with the fuzzy expert alerting system, so that the system will be able to issue early warnings based on predicted hydrological and/or meteorological parameters values. Furthermore, future study will explore the possibility to construct prediction models for the other variables on shorter time-scales than the one-day ahead prediction. We also intend to investigate various energy-preservation policies and the trade-off between prediction accuracy and data quality, which will allow us to deploy the water quality monitoring system in aquasystems with limited sunlight. Finally, one of our future aims is the use of model-based reasoning for self-diagnosing the sensors of the LMSs and for predicting spatial pollution propagation among LMSs.

ACKNOWLEDGEMENTS

The project was co-funded by the European Social Fund & National Resources - EPEAEK II - ARCHIMIDIS.

REFERENCES

Bing, S.L., Bullister, J., Murray, J., Sonnerup, R., Anthropogenic chlorofluorocarbons in the Black Sea and the Sea of Marmara. *Deep Sea Research Part I: Oceanographic Research Papers*. 49 (5), 895-913, 2002.

Blockeed, H., Dzeroski, S., Crbovic, J., Simultaneous prediction of multiple chemical parameters of river water quality with tilde. *3rd European Conference on Principles of Data Mining and Knowledge Discovery*. LNAI 1704, Springer-Verlag, 1999.

Cheng, H., Yang, Z., Chan, C.W., An expert system for decision support of municipal water pollution control. Engineering Applications of Artificial Intelligence. Vol. 16, No. 2, Special Issue on "Applications of Artificial Intelligence for Management and Control of Pollution Minimization and Mitigation Processes", pp. 159-166, 2003.

Dzeroski, S., Demsar, D., Grbovic, J., Predicting chemical parameters of river water quality from bioindicator data. *Applied Intelligence*. 13, 7-17, 2000.

Fdez-Riverola, F., Corchado, J.M., CBR based system for forecasting red tides. *Knowledge-Based Systems*. 16, 321–328, 2003.

Hatzikos, E.V., A fully automated control network for monitoring polluted water elements. *4th International Conference on Technology and Automation*. Thessaloniki, Greece, 443-448, 2002.

Hatzikos, E.V., Bassiliades, N., Asmanis, L., Vlahavas, I., Monitoring Water Quality through a Telematic Sensor Network and a Fuzzy Expert System. *Expert Systems*. 24 (3), 143-161, 2007.

Hatzikos, E.V., Hätönen, J., Bassiliades, N., Vlahavas, I., Fournou, E., Applying adaptive prediction to sea-water quality measurements. *Expert Systems with Applications*. 36, (3-II), 6773-6779, 2009.

Hatzikos, E.V., Tsoumakas, G., Tzanis, G., Bassiliades, N., Vlahavas, I., An Empirical Study of Sea Water Quality Prediction. *Knowledge-Based Systems*. 21 (6), 471-478, 2008.

Hendee, J.C., An expert system for marine environmental monitoring in the Florida Keys National Marine Sanctuary and Florida Bay. *2nd International Conference on Environmental Problems on Coastal Regions*. Cancun, Mexico, 57-66, 1998.

Lee, H.K., Oh, K.D., Park, D.H., Jung, J.H., Yoon, S.J., Fuzzy expert system to determine stream water quality classification from ecological information. *Water Science and Technology*. 36 (12), 199-206, 1997.

Lehmann, A., Rode, M., Long-term behaviour and cross-correlation water quality analysis of the river Elbe, Germany. *Water Research*. 35 (9), 2153-2160, 2001.

Reckhow, K., Water quality prediction and probability network models. *Canadian Journal of Fisheries and Aquatic Sciences*. 56, 1150-1158, 1999.

Saunders, M.C., Sullivan, T.J., Nash, B.L., Tonnessen, K.A., Miller, B.J., A knowledge-based approach for classifying lake water chemistry. *Knowledge-Based Systems*. 18, 47–54, 2005.

Ecosystem concept in assessment of running waters degradation

K. Brabec[a,b], **J. Jarkovský**[a,b], **L. Dušek**[a] **and I. Holoubek**[b]

[a] *Institute of Biostatistics and Analyses, Masaryk University, Kamenice 126/3, 625 00 Brno, Czech Republic (kbrabec@iba.muni.cz)*
[b] *Research Centre for Environmental Chemistry and Ecotoxicology, Masaryk University, Kamenice 126/3, 625 00 Brno, Czech Republic*

Abstract: In comparison with past monitoring programmes targeted predominantly to water quality and sediments contamination the methods recently implemented into water management in Europe increased consideration of physical structure and processes of aquatic ecosystems. Water Framework Directive (WFD) is a flagship creating legislative preconditions for implementation of recent scientific knowledge on aquatic ecology and environmental indicators into water management practice.
Ecosystem concept being applied in the assessment systems developed for running waters implements linkages of hierarchical structure of fluvial ecosystems, hydromorphological characteristics of habitats, channels, floodplains and catchments, typology and scale-dependance of antropogenic impacts and biological responses. Indicative values of aquatic communities are based on their sensitivity to stressor, deviation of community characteristics from reference conditions. These response measures – metrics are combined in multimetric system developed for each organism group (biological element sensu WFD) and each stream type.
Development and calibration of new assessment systems is also associated with application innovative methods of statistical analyses and information technologies. Trecking the biological response to diverse anthropogenic impacts is more sophisticated and complex. Predictive modelling of reference communities has been included in some countries. Indicative characteristics of aquatic taxa (traits) are periodically updated by new scientific results.
Assessment of ecological status required by WFD has features and targets for which would be benefitial methodological linking to risk assessment methods traditionally applied in toxicological studies. Both methodologies are calibrated based on intensity of stressor-response relationship, requires quantification of risks about the timing of detecting a change, likelihood of its reversibility, detection of tresholds and consequences of management actions restoring non-acceptable status.

Acknowledgements: The contribution was supported by the Czech Ministry of Education, Youth and Sport (MSMT 0021622412 INCHEMBIOL) and by projects EURO-LIMPACS (GOCE-CT-2003-505540) and SP2e75008 VALENCE.

Keywords: Rivers, Water Framework Directive; Assessment systems, Structure of aquatic ecosystems, Risk assessment

Methodological approaches in ecological risk assessment (EcoRA)

L. Dušek[a,b], **I. Holoubek**[a] **and J. Jarkovský**[a, b]

[a] *Research Centre for Environmental Chemistry and Ecotoxicology, Faculty of Science, Masaryk University, Kamenice 3, 634 00 Brno, Czech Republic (dusek@iba.muni.cz)*

[b] *Institute of Biostatistics and Analyse, Masaryk University, Kamenice 3, 634 00 Brno, Czech Republic*

Abstract: Environmental (Ecological) Risk Assessment (EcoRA) is a complex process of assessing the influence of human activities on ecosystem duality and vice versa. The endpoint of EcoRA is a prospective or retrospective assessment of stress factors influence (chemical contaminants, anthropogenic interventions or natural disasters) on ecosystems and their parts. The ecological risk assessment is a rapidly evolving discipline that is quickly being incorporated into the daily routines of industry and government. The EcoRA framework has recently been completed or is under development in numerous countries of Europe, in USA and Canada. In present time, many specific approaches and models exist in numerous sub-disciplines of EcoHRA, , such as chemical-property and fate estimation, toxicity, biological uptake, and population effects with limited co-ordination across these sub-disciplines. Recent development of information and communication technologies provides very powerful platforms for effective processing of multiple data sources. Particularly, web searches, database systems and data mining tools oriented on key environmental components and their descriptors, regionally-specific data aggregation, mapping of exposure and segmentation of the region of interest using Geographic Information Systems technology, automated processing of laboratory tests, namely dose-response curves using eco-toxicologically relevant models, algorithms and statistical packages and probabilistic estimates of risk level and associated uncertainties and their reflection in standardization of used environmental information systems by EEA, EPA, OECD and UNEP.

Keywords: Ecological risk assessment, Environmental informatics.

1. RISK ASSESSMENT STUDIES AND THEIR METHODOLOGY

Environmental and human risk assessment can be characterized from the viewpoint of informatics as complicated processing of heterogeneous data (mostly retrospectively collected from various sources) leading to probabilistic estimation of some uncertain (prospective approach) or on the other hand relatively certain (retrospective approach) risk event. Key methodical steps of the whole process are summarized in Fig. 1 and can be simply defined as follows.

1. Problem formulation and hazard identification - Introduction to any reasonably designed study. It includes recognition of the area of interest, collection and aggregation of required information and preliminary focus on identified principal pollutants (stressors), source of contamination and most vulnerable environmental components and biological receptors.

2. Multi-component exposure assessment - exploration, identification and quantification of important exposure pathways. It includes modelling and summaries of accessible data as well as empirical estimates of environmental concentrations of proposed key pollutants.

3. Biological effect evaluation - Empirical phase focused on concentration-related or dose-related reactions of biological systems. Principal aim is to get parametric measures that identify biologically dangerous concentration levels. The process should not be limited only to laboratory testing, it works with ecosystem monitoring as well. Whenever we have access to regional or national bio-monitoring network we should use this data as very powerful information background.

4. Risk characterization - is completely computational process that leads to the probabilistic estimate of the risk. In fact it is stochastic aggregation of data from all the preceding methodical blocks.

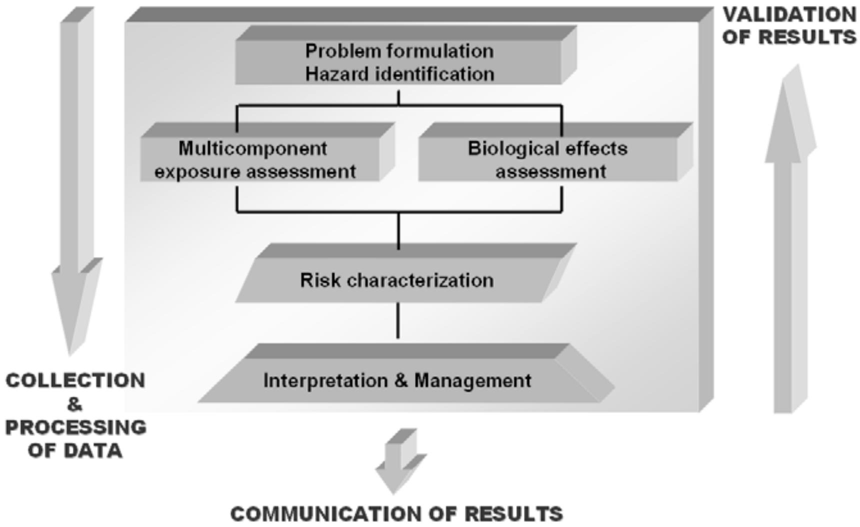


Figure 1. Environmental risk assessment (EcoRA) and key methodical steps

In other words, there are many inputs required and only limited number of outputs provided, however always with serious impact. The whole process can be also visualized as a circle that takes very different data and aggregates them according to given rules (Fig. 2).

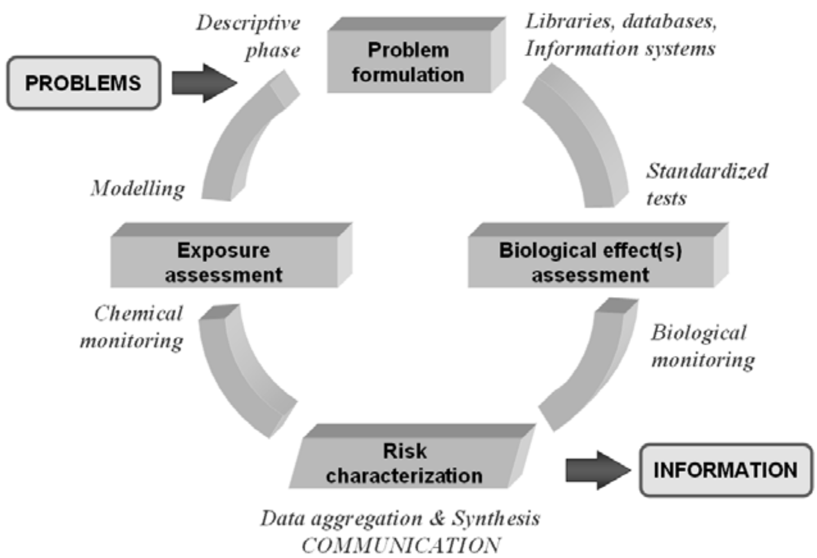


Figure 2. EcoRA methodology visualized as circle

The circle is constrained to rotate due to never–ending chain of problems entering the system. The image of „circle process“ is functional because not all phases must be necessarily employed in any type of problems. Different situations give to different phases very different weights. For example, the situation is thoroughly different if someone needs retrospective or prospective assessment than if it is sudden catastrophic situation like industrial accident or flood where we must in first line mitigate the immediate effects.

2. TIERED ANALYSES IN ENVIRONMENTAL RISK ASSESSMENT STUDIES

Heterogeneity of input data ranging from controlled laboratory bio-tests to multilevel time- or space-aggregated data implies heterogeneity of applicable experimental approaches. It is often very difficult to select the best approach and many later steps assume some specific result or outcome from the preceding measurements. That is why the experimental plans must be substantially more flexible as we can compare with some other types of research, for example with clinical trials. We must solve the complexity of the system, where no unique, definitely the best model can be recommended. We can of course move on according to plan, but from the first step we must categorize the problem - and here rather multiple (independent) lines of evidence are better than any single approach. This simply minimizes the probability of missing or omitting of some important facts and makes the solution sufficiently robust.

We adopt simple methodology from the first step and re-organize it in so called tiers as it is depicted in the scheme of Fig. 3. The consecutive tiers are gradually unravelled according to methodology, and repeated if the uncertainty is too high. As an example, we can perform some screening biological test to survey and point out the sites of interest, but then – in testing according to given scenario - similar tests can be repeated to verify the findings or to prove it in the other sites. Tiered style allows us to control the process and to react flexibly on unexpected changes or situations.

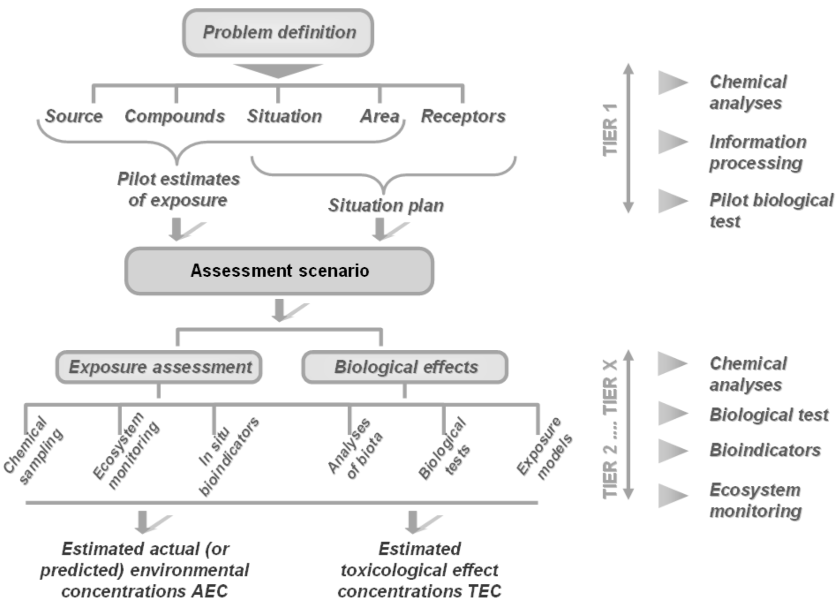


Figure 3. Tiered style of EcoRA methodology

3. INFORMATION AND COMMUNICATION TECHNOLOGIES IN ENVIRONMENTAL AND HUMAN RISK ASSESSMENT

Information and communication technologies (ICTs) are incorporated in each methodical step of formal risk assessment, from data gathering and analyses, through risk estimation to final validation of results and minimization of uncertainties.

Fig. 4 brings most important methodical streams with localization in EcoRA information flow. ICTs assist in building of assessment scenario and are indispensable in optimizing of experimental design. To summarize, information technologies and data analysis give the assessment process all its desirable properties:

- evidence-based background;
- sufficient information power;
- credibility and confidence;
- effectiveness and presentation skills.

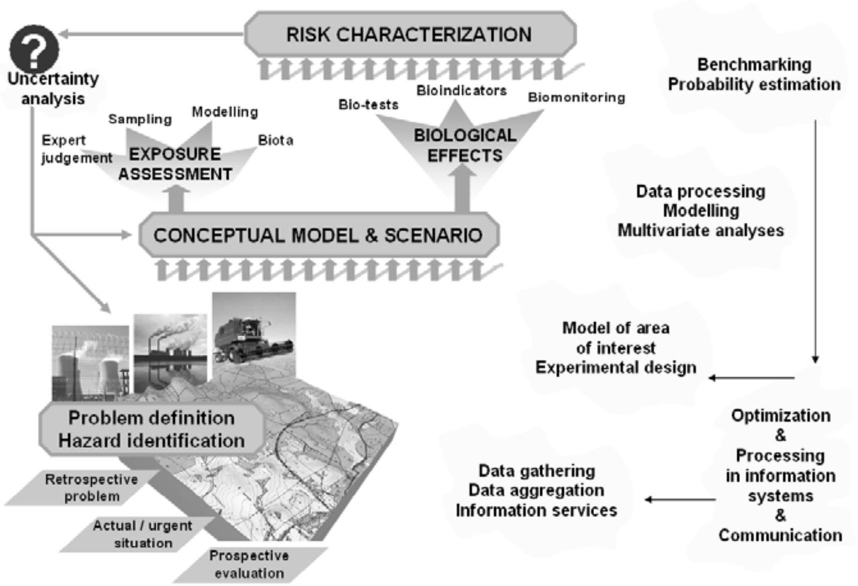


Figure 4. ICTs as indispensable part of each methodical step

Some important methodical tools and gains are summarized in the following paragraphs:

Evidence-based methodology. The "computational black box" of risk assessment covers a wide heterogeneity of input data ranges from controlled laboratory bio-tests to multilevel ecosystem descriptors measured repeatedly both in time and space. Although heterogeneity of inputs could imply heterogeneity of applicable computational approaches, one of the substantial roles of informatics is to standardize the aggregation and subsequent analyses of data. We must accept data management rules as inherent part of the methodology that brings "evidence – based" approach and results.

Standardized data management. The evidence-based methodology is based on prospectively planned sampling and experimental design, with statistical reasoning of sample size and with objective targeting in the area of interest. The description of risk situation should be based on representatively described environmental components and all potentially influential stress factors. The parametric structure of endpoints is preset as representative for important biological receptors and absolute measures are evaluated with respect to objective benchmarks. Current informatics and statistics operate with technologies and algorithms that are able to standardize multi-level process like environmental risk assessment. Development of standardized and automated tools for data management is of increasing importance also due to strong impact on decision making process that follows the risk assessment.

Computational methodology and data processing. The principal role of informatics in environmental risk assessment cannot be narrowed only to standardized gathering and aggregation of primary data. All phases of the process (hazard identification, exposure assessment, dose-response monitoring) are intrinsically associated with some level of uncertainty and so the final conclusions are based on stochastic analytic methods. At this point we must accentuate key role of GIS technology, multivariate processing of environmental bioindicators and finally very important dose-response modelling and probabilistic characterization of risk. Each of these methods represent unique field of computational science with its own background. Notwithstanding the methodical variety, environmental risk assessment must assimilate only verified approaches with sufficiently robust algorithms, suitable even to heterogeneous or incomplete data. Reliable computational methodology bridges the gap between environmental and experimental data and makes the whole process as effective as it is possible.

Flexibility to incorporate new parameters and technologies. Similarly like other biological sciences, current development in ecotoxicology is attacked by occurring high-throughput technologies that accelerate toxicity testing and push forward our understanding of mechanisms of eco-toxicity. The seamy side of the development is overproduction of experimental data that are typically measured in artificial conditions and under influence of environmentally irrelevant doses of stress factors. The role of informatics is to incorporate this novel experimental dimension in the frame of routine environmental assessment. The applications however cannot be blind and must stand on proper dose-response measurements treated with respect to inherent uncertainty (extrapolation from artificial conditions, inter-species variability, concentration levels …). As an example we could mention choices that are relatively very easy when dealing with high-dose exposures but that become substantially more difficult in low-dose range of exposures. Another example may be the advantage we could take from fascinating development in the field of genomics and budding toxicogenomics that can contribute to environmental risk assessment of carcinogenic compounds. Computational science indeed is the only change how to open the door for these completely new parameters without violation of standard routines.

Development of information systems. Very important contribution of informatics to environmental risk assessment that can never be neglected when data is to be automatically delivered directly to the point of decision. The systems substantially enhance individual susceptibility into risk assessment and ensure rational comparative evaluation or benchmarking in networks of differently experienced users. Only accessible information can form the basis for decision making and only sophisticated information systems can guarantee accessibility of information from large (bio)monitoring programs. Informatics viewed in terms of information systems also brings important training and educational platform.

Expert systems. The last aspect that should be mentioned in the context of environmental informatics is expert information service. Hundreds of environmental monitoring projects had been performed (and currently are carried out) in tens of countries and indeed it seems that there is even redundancy of environmental data at least in our European region. The critical problem however is availability and correct interpretation of the data. Outputs of governmental and non-governmental projects typically exist in separated databases of many institutions and their merged browsing and analyses are complicated and require participation of specialists from many fields. This time-consuming process often strongly limits the availability of reasonable data at the point of strategic need. Namely data on environmental fate of persistent organic pollutants requires long-term analyses and are difficult to interpret without multiple entries. All these problems can be technically solved through safe expert systems that aggregate relevant data sources with final automated analytic tools. This is the future of “responsible environmental reporting”.

Writing about ICTs in the EcoRA methodology, we should never forget communication of the results. Risk estimate itself is complicated end-point, conditioned by many items and circumstances. Risk should be communicated as probability of some hazard event that however can reach different levels for different biological systems. Without adequate presentation, there is a danger of incorrect interpretation and the examined situation can be unacceptably underestimated or omitted. ICTs have the power to emphasise the right findings to the right people. We can make the results accessible even for general public or to provide interactive tools for the evaluation of different scenarios. There are numerous target groups for the communication of EcoRA results with very different requirements on the form and content of EcoRA reporting (Fig. 5). Only modern communication technologies including open public web portals can serve for such complicated communication field.

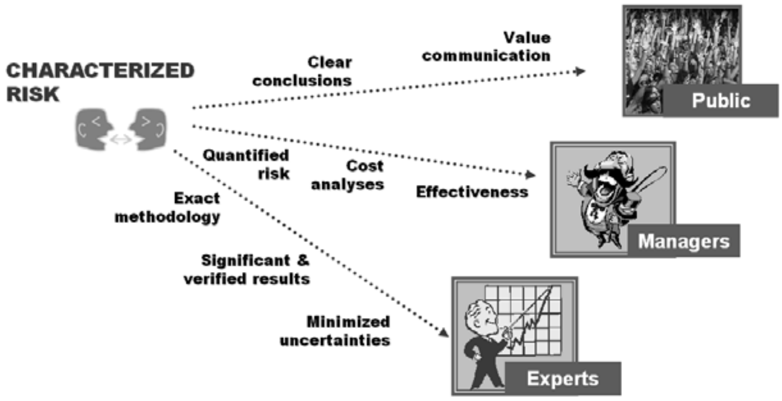


Figure 5. Communication demands on information service from EcoRA

4. MORE PRAGMATIC VIEW INTO THE COMPLICATED ECORA METHODOLOGY

Although the methodology seems to be relatively complicated and rigid, it is not true. It really has to be easily adapted for the assessment of many types of ecosystems like soil, surface fresh waters or some special type of urban area. In practical realization, each methodical block is responsible for answering of some principal questions and provides estimates with minimized uncertainties and risk of bias. Following paragraphs summarizes such simplified, question-oriented approach.

Any risk assessment starts with problem definition in several prominent components. First, we must know the situation where we are and at least approximately what has happened and what is required. Second, we must, at least hypothetically, identify potential source or problem. Once we know these two components we should concentrate on the following key questions that are to be answered, as based on the assignment:

- Are we able to prioritize further methodical steps?
- Are we able to define sampling and measurement plan?
- Are we able to stratify the area of interest according to risk?
- Are we able to recognize risk situation?

These questions target the scenario of further assessment. The answers of course require sufficient awareness of the situation in the area of interest that can be get from accessible databases and libraries or from pilot screening survey (screening of exposure pathways, pilot investigation of potential sources of toxic compounds, etc.).

That is why the process of assessment is unavoidably initiated by rather inventive searching for information where we typically work with three types of data sources (Fig. 6):

- widely accessible databases of toxic compounds and their characteristics,
- local databases and standards,
- pilot results of some tests and ad hoc measurements.

Gained information must be processed and the sources must be prioritized, filtered and then aggregated in order to get really usable input into assessment scenario. The information processing is typically projected into the area of interest and so most recommended tool for this phase is some type of geographical information system (GIS).

The assessor must be aware of relative spatial and temporal scales of the hazardous situation being assessed. The segmentation of the area of interest determines the extent and range of exposure and identifies the sites with increased probability of risk attack. However, we must take into consideration the assignment for the assessment itself, it means to correlate between measurement feasibility and risk situation and assessment interest that is placed into by experts, public, environmental activists, owners or managers as it is indicated in Fig. 7.

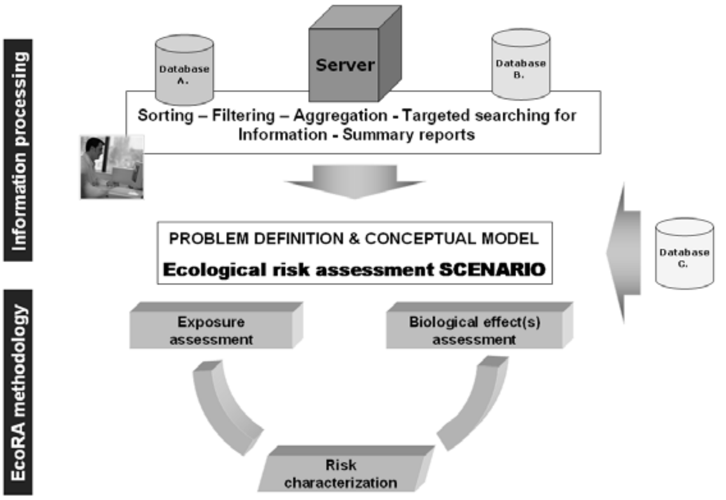


Figure 6. Problem definition and applicable data sources

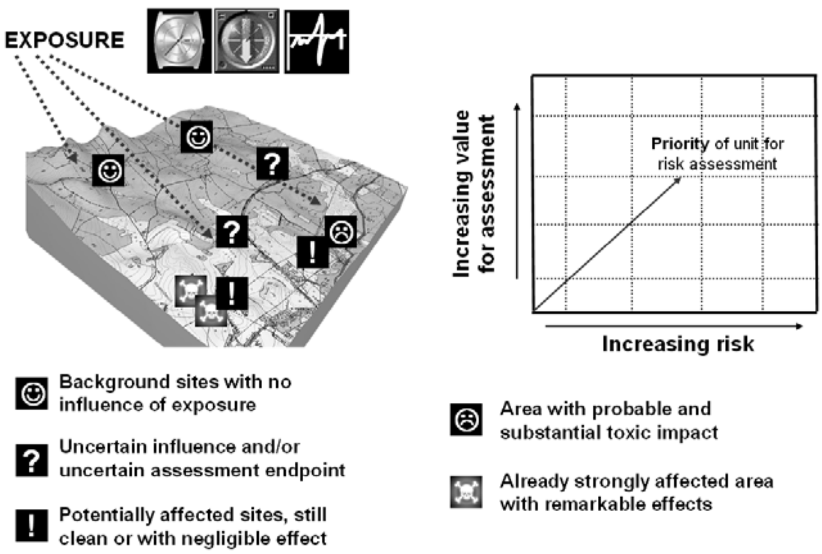


Figure 7. Problem definition generates comprehensive situation plan

The situation plan thus opens following important questions that can be answered only in comprehensive experimental plan:

- How to target heterogeneous area of interest? And what is optimal and safe sampling plan for in situ investigation?
- What is the optimal set of further analyses?
- How to effectively combine ecological and chemical measures and laboratory tests?

The assessment scenario can be regarded as milestone of the assessment process, because it definitely summarizes obtained information and generates hypotheses and strict plan of further movement. Fig. 8 depicts basic principle of categorization of the scenarios working with primary and secondary exposure pathways. The idea of this approach is simple – most of the further measurements are determined by the type of exposed environmental matrices. If we are able to distinguish contaminated matrices, it simply places the necessary applied tests. The approach also implies primary and secondary risk for prospective studies. Distinguished separation of different types of ecosystems (matrices) can also be easily associated with legislation and guidelines valid for the field.

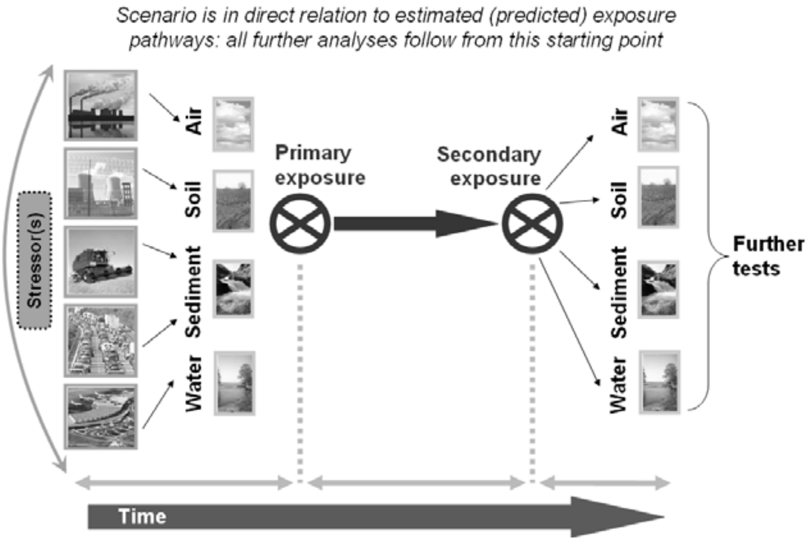


Figure 8. Assessment scenario and basic principle: „Where is the problem“

Here we must strongly emphasize very important dimension of any practically targeted risk assessment scenario, it means effectiveness and pragmatic simplicity. EcoRa is very laborious work and so there is only limited space for research activities. We cannot rely on scientific and progressive data trying or even hoping that it improves our understanding common ecological data. Instead of that, the simplest and already verified set of methods should be applied in first line and additional analyses should be applied only in the case of remarkable uncertainties.

The selection of proper biological end-points is the last important step that should be prescribed in the scenario. This step is extremely important because the risk estimate cannot rely only on chemical data:

- biological tests can reveal some new effects that cannot be predicted from chemical data (this holds namely for heterogeneous environmental mixtures of chemicals)
- bioindication or monitoring survey can indicate the effects in past, that cannot be detected by chemical analyses

There are basically two strategies in biological effects assessment (Fig. 9): (1) model testing under defined conditions (mostly laboratory tests and bioassays) and (2) monitoring of real ecosystems that is often referred to as bioindication. Typically we employ different tests or tests using different biological models (receptors) to cover the heterogeneity and various food strategies of biological systems in the field.

Multiple biological tests of course require a wide spectrum of analytical methods as it will be explained in next paragraphs. At this point we should only emphasize, that outcomes of both chemical measurements (mostly applied during area recognition and exposure assessment – see Fig. 1, 2) and biological tests provide principally the same numerical output, concentration measure:

- environmental concentration (mostly obtained on the basis of chemical monitoring) as estimate of the real (or predicted) contamination of target matrices
- biologically safe concentration (obtained from laboratory tests, biomonitoring and/or in situ assays) as the estimate of the concentration that does not cause biologically harmful effect or on the other hand, the concentration that causes some degree of the effect.

Both types of concentration estimates are then mutually related in probabilistic models or ratios to estimate the probability of risk event (Fig. 10). We typically mutually relate two or more probabilistic distributional models and try to find concentration cut-off point that determines the highest acceptable environmental concentration that is still biologically safe.

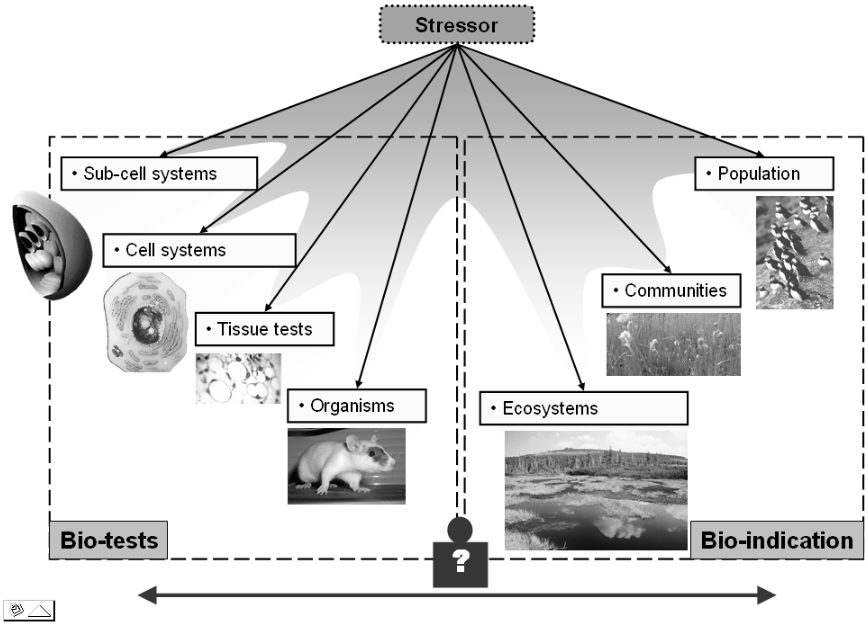


Figure 9. Two basic strategies for the selection of biological end-point

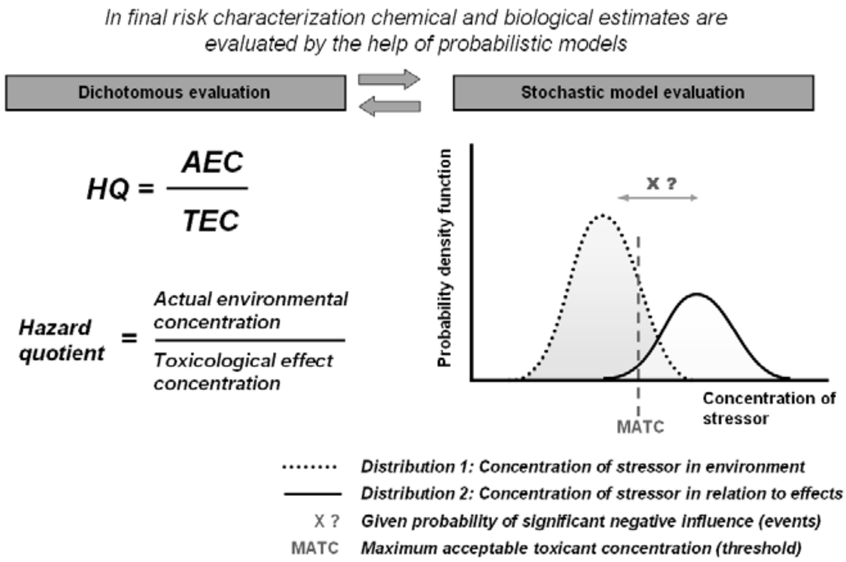


Figure 10. Simplified example that demonstrates risk estimate

5. MOST IMPORTANT ASPECTS OF ANALYSIS OF ENVIRONMENTAL DATA AND RELATED RESEARCH TOPICS

Data analysis plays strategic role namely in the following fields:

- Data models, auditing and validation of information sources. Relevant arrangement of input information, description of the situation, hazard identification.
- Regionally specific aggregation of accessible data, quality control and GIS models of the area of interest, mapping of exposure pathways and levels.
- Hierarchical structure and prioritization of biological indicators according to ecological criteria, susceptibility to stress factors and accessibility for measurements. Application of data from large-scale biomonitoring networks. Searching for biological "hot spots" and reference standards for biological systems in the area of interest. Selection of proper biological receptors. Currently, more and more sophisticated techniques are being applied in this field, namely data mining technology.

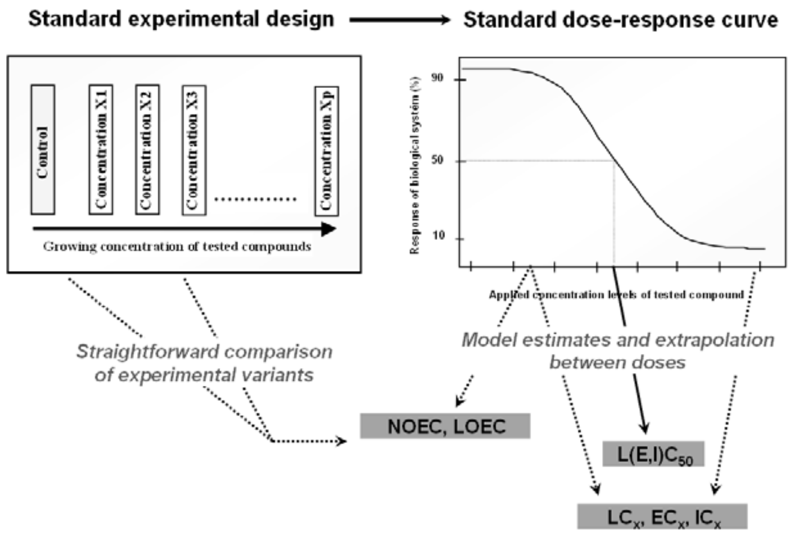


Figure 11. Standard output of bio-tests: dose-response curve

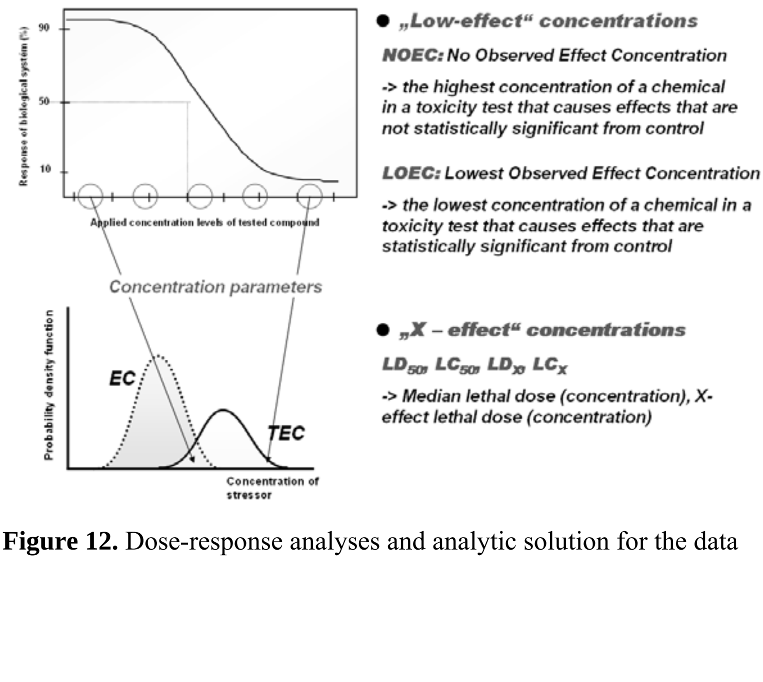


Figure 12. Dose-response analyses and analytic solution for the data

- Standardization of laboratory tests and their outputs (i.e. concentration measures that are related to some level of biological effects: NOEC, LOEC, NOAEL, LOAEL, IDx, LDx, EDx, ICx, LCx, ECx). Practical implementation of these outputs in decision making includes following topics (see also examples in Fig. 11-13):
  - standardized and reliably estimated concentration measures
  - estimation and formal mathematical description of different types of dose-response curves (because the shape of the relationship is important as well)
  - typology of dose-response curves for different situations and stressors
  - mathematical processing of complicated, but real dose-response patterns (see Fig. 13).
- Benchmarking of ecosystem abiotic and biotic characteristics, risk characterization and regionally specific interpretation. This approach generates more complicated data structure than laboratory biotests with the following specifics:
  - It is impractical to measure these complex parameters in predictive design, where the conclusion must be made before any large-scale release and effects can occur, the evaluation is then retrospective not prospective.
  - In fact it is virtual reconstruction of the conditions of the injured environment prior to the pollution release so that it can be compared to the injured condition.
  - In the case of such complex parameters, standard statistical methodology fails or we have no time to verify hypotheses in repeated experiments.
  - An alternative methodology to common statistical scoring must be used and it is multivariate statistical modelling: a) processing of so called calibration data set that is representative for the ecosystem or b) work with environmental gradient data. We rank hazardous and clean sites using a linear discrimination with the support of training set of well characterized sites that are assigned to categories on a scale hazard. The site attributes are used to establish a discriminant function. If the function separates the categories sufficiently, new sites are classified on the hazard scale by applying the discriminant function (see Fig. 14, 15).

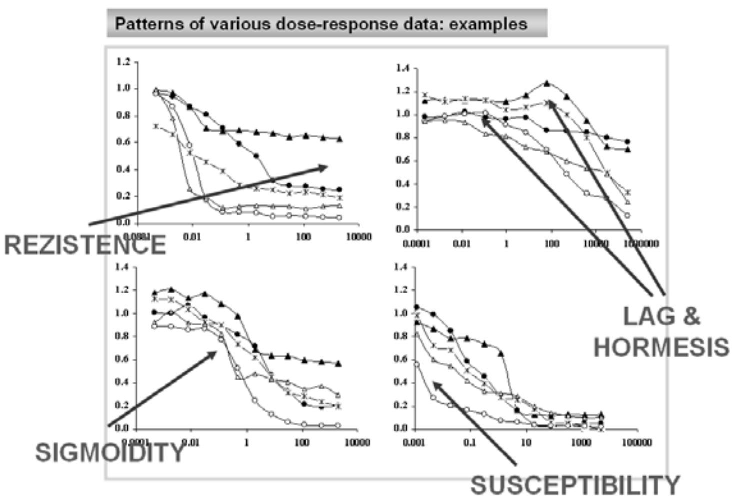


Figure 13. „Real“ variability of dose-response curves and related parameters

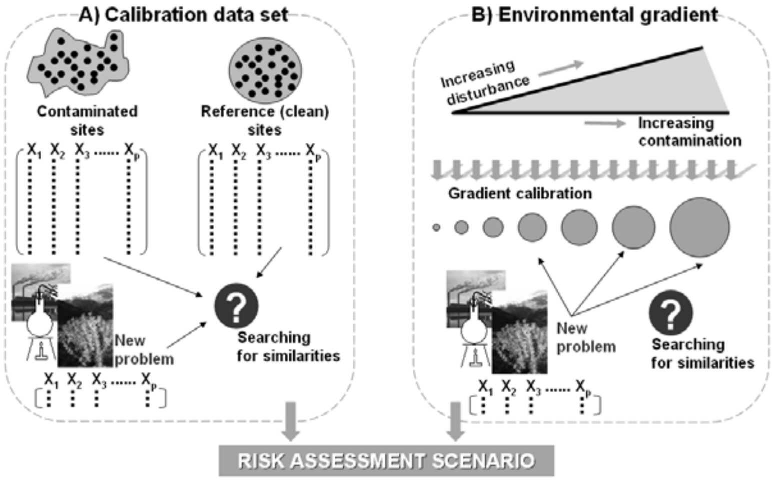


Figure 14. Scheme of model processing of ecosystem data

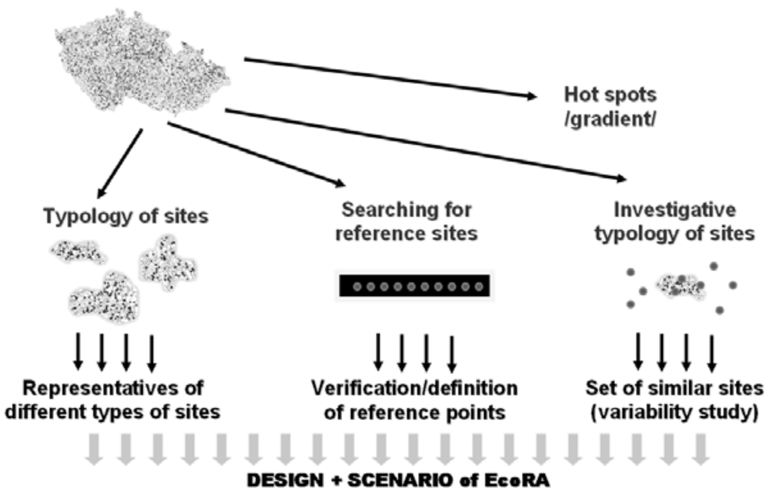


Figure 15. Modes of multivariate processing of ecosystem data

6. MOST IMPORTANT ASPECTS OF ANALYSIS OF ENVIRONMENTAL DATA AND RELATED RESEARCH TOPICS

Computerized environmental information systems are rapidly being rolled out into use, but typically they are not complex enough to cover all attributes describing ecosystem level of organization. Lack of available software tools or algorithms mostly leads to simplified evaluation of some components or to asymmetric assessment based more on easily accessible laboratory data. Evaluation of computational and expert capability of environmental information systems lags far behind their application and mostly we know very little about their impact on final estimation of risk. The main ambition of informatics in this field is then to solve the following problems that are commonly associated with or sometimes even generated by different types of information systems:

- how the information systems affect methodology of gathering of primary data
- how to effectively optimize collection, aggregation and filtering of heterogeneous, large scale data (multidimensional systems, grid computational techniques, data mining, GIS methodology)
- how to develop user-oriented expert systems ensuring validated and flexible feedback of the assessor towards complexity and quality of primary data

- how to optimize information systems for monitoring of persistent organic pollutants (POPs) (data needs, modelling of persistence, multimedia models, expert services evaluating long-range transport of POPs, specific biological effects, etc.)
- how to maximize ration between costs and benefits in computer-assisted risk evaluation
- how to implement information systems in the following important, but complicated fields (data models, computational and technological requirements):
    - processing, aggregation and evaluation of biotic data at population or ecosystem level
    - centralized procurement of large scaled data, retrospective analyses of environmental time series
    - processing of biodiversity data in environmental risk assessment
    - generation and simulation of different scenarios (exposure pathways, toxic effects, ...)

Some of the barriers that limit complexity of environmental information systems are obvious. In contrast to relatively easy issue of novel bio-test or indication method, timescales are long and the ability to switch a system to more complex level is limited by cost and organizational constraints. In some key fields (biodiversity monitoring, fate of persistent compounds, population epidemiological risk) it is hard to collect representative data in relatively short period of time. Processing and interpretation of often incomplete data from natural systems cannot be straightforward and is dependent on effective computational techniques. It is mostly not possible to carry out adequate assessment of the large scale systems with techniques that have been successful in smaller systems with limited heterogeneity.

7. CONCLUSION

We hope that this short overview will help to reduce confusion and sometimes conflict over the roles of information technology in environmental and human risk assessment, and a lack of scientific as opposed to a managerial simplified approach in the development of expert systems. Many of the mentioned problems are in detail characterized in the next chapters of the proceedings.

ACKNOWLEDGEMENTS

The project was supported by the Czech Ministry of Education, Youth and Sport (MSMT 0021622412 INCHEMBIOL).

REFERENCES

All relevant citations on the development of EcoRA methodology can be found in the research report of the IDRIS project: Holoubek, L. Dušek, M. Machala, P. Čupr, K. Bláha: Project IDRIS - Ecological Risk Assessment - regional approaches. In: Assessment and Management of Environmental Risks: Methods and Applications in Eastern European and Developing Countries. Kluwer Academic Publishers, 283-298, 2001.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pilmmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Experiences and case studies analysis on risk-based Decision Support Systems for contaminated water and land management

S. Gottardo[a], P. Agostini[a], A. Marcomini[a], A. Critto[a], G.W. Suter[b] II

[a] *Department of Environmental Sciences, University Ca' Foscari of Venice, Calle Larga Santa Marta 2137, 30123 Venice, Italy (marcom@unive.it)*

[b] *National Center for Environmental Assessment, U.S. Environmental Protection Agency, Cincinnati, Ohio OH 45268, USA (suter.glenn@epamail.epa.gov)*

Abstract: Nowadays both freshwater and soil ecosystems in Europe are often highly contaminated by mixtures of toxicants thus making their rehabilitation and protection very complex. In order to properly address different environmental issues and effectively manage contaminated sites through specific interventions a large amount of data and information needs to be collected and processed by competent authorities and industries in Member States. Moreover the assessment and evaluation of environmental risks is widely carried out and strongly recommended in most European directives and regulations, e.g. in the frame of the Water Framework Directive 2000/60/CE where it is consolidated and in the frame of the Soil Thematic Strategy where it is still debated. The results obtained are used to address the decisional process and the identified risks (such as affected assessment endpoints, areas of major concern, likely stress factors) and should be communicated to any stakeholders and to the public in a transparent way. In this context tools that facilitate storage and organization of data provided by various information sources as well as analysis of complex environmental issues are of great benefit. Several research and application activities have been undertaken in Europe and in USA during recent decades to develop dedicated tools supporting data management, performance of quantitative analyses, evaluation of simulated alternative scenarios, participatory processes and results visualization. Efforts have been particularly focused on Decision Support Systems (DSS), which we define as computer-based tools that integrate and evaluate economic, social and environmental information through multiple functionalities in the light of stakeholders' preferences and in a sustainable use and development perspective. Several methodologies and tools can be implemented within a DSS such as Ecological Risk Assessment (ERA), Multi Criteria Decision Analysis, Spatial Analysis and GIS, environmental quality indicators and indices. A wide array of DSSs including case study applications and best practises for the risk-based assessment and management of contaminated land and inland waters in Europe and US are presented. The overview compares different types of DSS based on legislative framework, scale of analysis, functionalities, methodological and structural elements, technical and web-based features and flexibility thus allowing identify main gaps and challenges faced by both users and developers. Specifically, dedicated Ecological Risk Assessment methodologies implemented in DSS for risk estimation or causal analysis are presented. Relevant issues and needs are then addressed and discussed in the specific case studies. For example a risk-based DSS for EU Water Framework Directive (WFD)-compliant river basins assessment and management is under-development within the EU project MODELKEY (Contract-No. 511237 (GOCE)). It implements an integrated risk assessment methodology based on a Weight of Evidence approach with a Fuzzy Inference System for environmental quality evaluation of fluvial ecosystems. Moreover a prioritization procedure for identifying hot spots along the river basin through integration of environmental and socio-economic information is included. Another ERA-based tool dealing with causal analysis of biological impairment detected in aquatic environments is the Causal Analysis/Diagnosis Decision Information System (CADDIS). It aims to improve the assessment of causes of biological effects, by providing a formal inferential methodology and technical content useful for implementing the method. Although the primary application of CADDIS has been to freshwater systems, it is based on principles that are applicable to any ecosystem, and it has recently been applied to contaminated terrestrial systems. The DSS-ERAMANIA was developed specifically for terrestrial ecosystems. It is based on a Triad approach and includes two modules: "Comparative Tables" and "Integrated Ecological Risk Indexes", both based on Multi Criteria Decision Analysis (MCDA). The first one compares different measurement endpoints (i.e. bioavailability tools, toxicity tests and ecological observations) to guide the expert/decision maker in the choice of the suitable set of tests to be applied to the case study. The second module provides qualitative and quantitative tools allowing the assessment of terrestrial ecosystem impairment (i.e. the impairment occurring on biodiversity and functional diversity of the terrestrial ecosystem) by integrating complementary information obtained by the application of the measurement endpoints selected in the first module. Case study experiences and related tools are collected together in the book "Decision Support Systems for Risk Based Management of Contaminated Sites" edited by Marcomini, Suter and Critto (Marcomini et al., 2009).

Keywords: decision support systems, risk assessment, contaminated site management, environmental quality, freshwater and soil ecosystems.

1. INTRODUCTION

Many factors comprise the decision-making context for the management of contaminated land and water. Assessment of those factors (e.g., environmental, socio-economic, technical, spatial, legislative) requires multiple approaches including ecological and human health risk assessment, comparative risk assessment, remedial investigation/feasibility study, economic assessments, site characterization, and integration models. Decision Support Systems (DSSs) for the management of human health risk assessment are increasingly used both in Europe and USA and those for ecological risks are gaining attention. A DSS is a computer-based tool that addresses integration and evaluation of economic, social and environmental information through multiple functionalities in the light of stakeholders' preferences and in a sustainable use and development perspective (Loucks, 1995; Shim et al., 2002).

A recently published book entitled "Decision Support Systems for Risk-Based Management of Contaminated Sites" describes the context of using DSS and provides a collection of several tools that are currently available in Europe and USA (Marcomini et al, 2009). In this paper, the focus will be on tools addressing ecological risk assessment, briefly presenting their characteristics and application examples. Additional details about the tools can be found in the relevant chapters of the book.

2. SOME US AND EUROPEAN DSSs FOR ECOLOGICAL RISK ASSESSMENT

The general description of tools concerned with ecological risk assessment is reported in Table 1. MODELKEY, ERA-MANIA and DITTY were developed in Europe, while CADDIS, SADA and BASINS are products of USA institutions.

The tools have some common aspects, particularly as far as methodologies are concerned, but also peculiarities. For examples, the DSS-ERAMANIA is a decision support system implementing a site-specific Ecological Risk Assessment (ERA) procedure (US-EPA, 1998) and supporting the experts and the decision makers in the assessment of contaminated soils. It was developed according to a Triad approach (Long and Chapman, 1985), where the results provided by a set of measurement endpoints are evaluated and integrated to support the assessment and characterization of ecosystem impairment caused by the soil contamination.

Table 1. DSSs for ecological risk assessment. (WOE = Weight of Evidence, MCDA = Multi Criteria Decision Analysis; DPSIR = Driving Force, Pressure, State, Impact, Response; IRI = Integrated Risk Indices; GIS = Geographical Information System)

Tools	Scale of analysis	General description	Main methodology	Reference
ERA-MANIA	Site-specific Soil	It is composed of two modules: Module 1 ("Comparative Tables") aims at comparing the different measurement endpoints belonging to each LOE (i.e. bioavailability tools, toxicity tests and ecological observations) to guide the expert/decision maker in the choice of the suitable set of tests; Module 2 ("Integrated Ecological Risk Indexes") provides qualitative and quantitative tools allowing the assessment of terrestrial ecosystem impairment.	Triad approach WOE MCDA	Critto et al,. 2007 Semenzin et al., 2007
MODELKEY	River basin Aquatic systems	It aims at interlinking and integrating different analytical tools and exposure/effect models in order to evaluate risks posed by pollution to aquatic ecosystems at river basin scale and to identify areas (hot spots) in need of management. The tiered procedure for evaluating ecological risks is based on two phases accomplishing multiple functions at both river basin and site-specific scales: data exploration and evaluation, quality status classification, identification of causes, economic analysis of water uses, hot spot prioritization, and provision of monitoring recommendations. Environmental information is integrated with socio-economic factors related to different water uses in order to prioritize hot spots to be managed.	DPSIR framework IRI MCDA WOE	www.modelkey.org
DITTY	Coastal lagoons Aquatic systems	It provides a common and flexible framework to ease the integration of the outputs of different models and analyses, as well as to deal with the diversity of socio-economic and environmental characteristics of several application sites. The DSS helps to generate the alternatives for comparison. Mathematical and analytical models (e.g., biogeochemical, hydrodynamic, ecological, and socio-economic models, geographic information systems, etc.) are used to simulate the alternatives, and to provide corresponding system performance indicators related to the decision criteria. MCDA is finally applied to evaluate and rank the alternatives on the basis of both the values of the indicators and the interaction with the decision maker.	Alternatives generation MCDA	http://www.dittyproject.org/
CADDIS	River basin Aquatic systems	It is an on-line decision framework for identifying the stressors responsible for undesirable biological conditions in aquatic systems. It helps practitioners choose which causes to consider, based on sources, site information, and observed biological effects. A series of conceptual models illustrates connections between sources, stressors, and effects. Statistical tools are provided to analyse evidence. A guide to all of the types of evidence may lead practitioners to organize existing evidence and seek additional evidence. Inferential guidance helps users to identify the most likely cause.	Stressor Identification Scoring system	http://www.epa.gov/caddis
BASINS	River basin Aquatic systems	It is a DSS for multipurpose environmental analysis by regional, state, and local agencies performing watershed and water quality-based studies. BASINS integrates environmental data, analytical tools, and modelling programs to support development of cost-effective approaches to watershed management and environmental protection, making it possible to quickly assess large amounts of data in a format that is easy to use and understand. In particular, it combines aqueous contamination for waste sites with other sources to estimate combined risks and to predict the benefits of remediation.	Prediction models Generation of Scenarios (GenScn) GIS	http://www.epa.gov/waterscience/basins/
SADA	Site-specific Soil	The integration of the GIS with ecological risk assessment allows the spatial presentation of data, comparison of model results to ecological benchmarks, exposure modelling in a spatial context, and development of movement-based exposure models. These capabilities allow for easier and implementation of screening-level hazard quotient (HQ) approaches that are commonly performed for terrestrial ecological risk assessments.	Terrestrial ecological risk assessment procedures GIS	http://www.tiem.utk.edu/~sada/download.shtml

In the Triad approach, the measurements generate three Lines of Evidence (LoE): environmental chemistry, ecotoxicology and ecology, whose integration should pragmatically reduce the uncertainty in the risk estimation. For the DSS, a framework including three subsequent investigation levels (i.e. tiers) was implemented that enables the completion of the risk assessment once the provided answer is unequivocal as characterized by a relatively small uncertainty, ensuring at the same time an adequate financial investment.

MODELKEY DSS, developed to address decision-making processes in compliance with the EU Water Framework Directive (WFD) (EC, 2000) requirements, is a system characterized by an "open configuration" able to manage and integrate different types of data, parameters and models and freeing end-users to include their own specific tools. It implements an Integrated Risk Assessment methodology based on a Weight of Evidence (WoE) approach (Burton et al., 2000) and implementing a Fuzzy Inference System (von Altrock, 1995) for environmental quality evaluation of fluvial ecosystems. Moreover a prioritization procedure aiming at identifying hot spots along the river basin through integration of both environmental and socio-economic information is included. The MODELKEY DSS software system is implemented in an open-source GIS, the uDig platform.

CADDIS is an on-line decision framework developed in response to requirements under the U.S. Clean Water Act to develop plans for restoring impaired aquatic systems. In particular, it provides a means to determine the causes of observed biological impairments. A series of conceptual models illustrates connections between sources, stressors and effects. Other tools are used for analyzing data and interpreting results as causal evidence. By a Step-by-Step procedure, practitioners are guided through fifteen different types of evidence. A consistent system for scoring the evidence, adapted from Susser, facilitates the synthesis of the information into a final conclusion, through assignment of up to three plusses (+++) or minuses (---) for strongly supportive or extremely weakening evidence, respectively.

All together, some of the tools of Table 1 make use of GIS capabilities, others provide powerful analytical tools, and many include Multi-Criteria Decision Analysis (MCDA) tools to allow comparison and ranking of relevant information or management alternatives. They are developed to address either soil or watershed contamination, while DITTY is specifically devoted for coastal lagoons.

3. CASE STUDIES

Differences among tools and functionalities may be better evaluated by discussing their application. To this scope, Figure 1 reports some of the interface and graphics resulting from case study applications.

MODELKEY DSS has been tested by means of applications to the three case studies of the MODELKEY project, i.e. Elbe, Scheldt and Llobregat River Basins. End-users for each case study have evaluated the DSS functionalities, interfaces and outputs. Preliminary results have been obtained for both Llobregat and Scheldt river basins. The Llobregat has a typical Mediterranean regime and represents the second-longest river of the Catalonia Region (Spain). In coastal urban areas waste waters represent a significant input to the river (Kuster et al. 2008). It also receives industrial wastewaters as well as surface runoff from agricultural areas that leads to contamination of organic compounds (e.g. pharmaceuticals, estrogens and pesticides). The Scheldt (France, Belgium, The Netherlands) is one of the most complex international Pilot River Basins selected by the European Commission in order to test WFD implementation. Long stretches of river are canalized and a number of its tributaries are subjected to the tides. The Scheldt is a highly urbanized and heavily built-up area. Households and industry represent the two major driving forces of the changes in the water system. A set of biological indicators of the macroinvertebrates and phytobenthos communities were calculated using monitoring data collected in 2003-2004, such as Shannon (Shannon and Weaver, 1949), Margalef (Margalef, 1984), Biological Monitoring Working Party (Armitage et al., 1983) and SPEcies At Risk (Von de Ohe et al., 2007). In addition Potentially Affected Fraction (PAF) (Posthuma et al., 2002) and Toxic Units (TU) were applied as chemical indicators while BOD5, NO3, PO4, pH were used as physico-chemical indicators. Indicator values at various sampling stations were compared with reference sites available for different water body types and integrated through a dedicated procedure. Final results were visualized on GIS maps by means of pie-charts showing the sampling point's percentage of membership to WFD quality classes (i.e. high, good, moderate, poor, and bad). Moreover the end-user can go back to visualize intermediate results by means of AMOEBA (a general method for ecosystem description and assessment) yardsticks (Brink et al., 1991): this facilitates identification of the most affected biological communities (e.g. macroinvertebrates rather than fish) and the most responsible causes of impairment (e.g. eutrophication or toxic pressure) at the sampling site of interest.

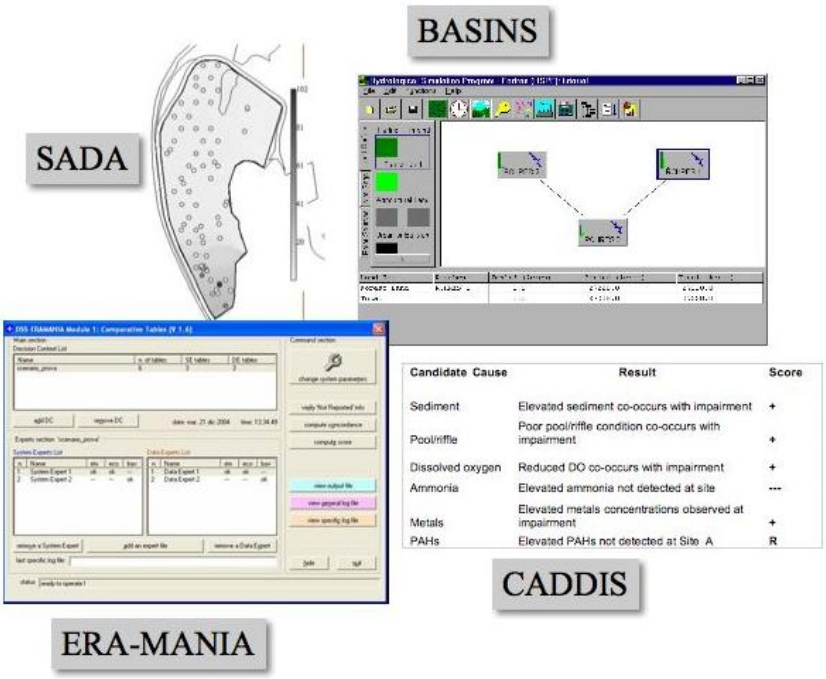


Figure 1– Interfaces and results of some of the tools

The DSS-ERAMANIA was preliminary applied to the Acna di Cengio contaminated site located in the Savona province (Italy) and extending over ca 550000 m2. From 1882 to 1999, production changed from dynamite and tri-nitrotoluene to lighting gas, chemical products (e.g. nitric acid, phenol, sulphuric acid), dyes and pigments, and finally to β-naphthol and phtalocyanine, causing relevant environmental problems. In December 1998, it was identified as one of 14 contaminated sites of national interest to be reclaimed. The preliminary application of DSS-ERAMANIA to the Acna contaminated site led to a transparent selection of measurement endpoints for each Triad LoE (i.e. Module 1 application), and to a comprehensive (quantitative and qualitative) evaluation of the impairment occurring on the site of concern (i.e. Module 2 application). Moreover, Module 1 turned out to be a useful tool for promoting discussions among experts to reach a common judgment on advantages and drawbacks of existing experimental tests. Module 2 confirmed its capability of bridging the gap between experimental test results on the one hand and site-specific risk estimation and evaluation for terrestrial ecosystems on the other hand. ERA-MANIA's easy interface guided the application of the system, as reported in the case of Module 1 included in Fig. 1.

Using CADDIS, the Little Scioto River case study assessed a 15-km reach of a river in north-central Ohio. Many point and non-point sources of pollutants are associated with the Little Scioto River, including a wastewater treatment plant and runoff from agricultural land uses and from the city of Marion, as well as from several contaminated industrial areas, including an abandoned wood treatment plant, a landfill, an appliance plant, and a rail facility. Three general geographical segments were separately considered—upper, middle, and lower—based on the biological conditions and causal analysis. The upper section was impaired by channelization and associated stressors, the middle section was extensively contaminated from a wood treatment site, and the lower part was affected by all of the upstream contaminants and metals. Figure 1 reports an example of the resulting scoring of the system.

SADA was applied to a K-770 Scrap Metal Yard less than 8 hectares in size and located on the western end of the Powerhouse Peninsula at East Tennessee Technology Park (ETTP) in Roane County, Tennessee. The yard operated during the 1940s as an oil storage area, and has operated since the 1960s as a scrap facility, although it is currently inactive. Tens of thousands of tons of metal were stored in piles at the site, but have since been removed. Most contamination at the site originated from the scrap piles. Considerable sampling has been conducted at this site for a full suite of analytes including metals and polychlorinated biphenyls (PCBs). SADA's spatial features allowed study of a selective remediation design by discretizing the site into a grid and identifying cleanup areas by remediating individual grid blocks in inverse order of risk magnitude (worst to least) until the cleanup objective is met, as visualized in Fig. 1.

The case study of Cottonwood Creek Watershed in Idaho County, Idaho, illustrates the use of BASINS as a decision support system for the needs of the US Clean Water Act and definition of the Total Maximum Daily Load. In particular, the Hydrological Simulation Program Fortran (HSPF) was used to represent the Cottonwood watershed's hydrology and fecal coliform loads to the creek. A spreadsheet was used to calculate bacteria related HSPF input parameters, and then the model was calibrated against fecal coliform monitoring data. The model was re-run with a percent reduction to nonpoint source loads to creeks, in addition to reductions in the cattle-in-stream parameter and faulty septic system "point sources" parameter, to determine how the state water quality standards could be achieved. Required nonpoint source load reductions ranged from 23% to 88%, when cattle-in-stream parameter and faulty septic system loads were reduced by 80-100%. In Fig. 1 the interface of the HSPF model is presented.

The case studies here briefly show how the different DSSs may provide analytical support for dissimilar conditions and management objectives.

4. CONCLUSIONS

This paper presents some of the tools available for ecological risk assessment, and their application to different case studies, as reported in the book by Marcomini and colleagues. In the same book, some unresolved issues and additional challenges for DSSs developers are also identified. Among them are, the flexibility to allow easy use of a DSS in various contexts, ease in changing input parameters or adding new models and functionalities; balancing the convenience and consistency of a relatively automated system against the transparency and flexibility of user-specified systems; and the effective involvement of end-users in the DSS development in order to avoid the common perception of these tools as black boxes.

ACKNOWLEDGEMENTS

The authors thank all the co-authors of the book "Decision Support Systems for Risk Based Management of Contaminated Sites" for their valuable contribution.

REFERENCES

Armitage, P.D., Moss, D., Wright, J.F., Furse, M.Y. The performance of a new biological water quality score system based on macroinvertebrates over a wide range of unpolluted running waters, Water research, 17, 333-343, 1983.

Burton, G.A., Chapman, P.M., Smith, E.P., Weight-of-Evidence approaches for assessing ecosystem impairment, Human and Ecological Risk Assessment, 8 (7), 1657-1673, 2002.

Critto A., Torresan S., Semenzin E., Giove S., Mesman M., Schouten A.J., Rutgers M., Marcomini A. Development of a site-specific Ecological Risk Assessment for contaminated sites: Part I. A multi-criteria based system for the selection of ecotoxicological tests and ecological observations. Science of the Total Environment, 379, 16-33, 2007.

EC (European Commission). Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000. Establishing a framework for Community action in the field of water policy. Official Journal of the European Communities, L 327-22.12.2000:72, 2000.

Kuster, M., Lo´pez de Alda, M.J., Hernando, M.D., Petrovic, M., Alonso J.M., Barcelo, D. Analysis and occurrence of pharmaceuticals, estrogens, progestogens and polar pesticides in sewage treatment plant effluents, river water and drinking water in the Llobregat river basin (Barcelona, Spain), Journal of Hydrology, 358, 112-123, 2008.

Long, E. R. and Chapman, P. M., 1985. A sediment quality triad: measures of sediment contamination, toxicity and infaunal community composition in Puget Sound. Marine Pollution Bulletin 16: 405-415.

Loucks, D.P., Developing and implementing decision support systems: a critique and a challenge. Water Resources Bulletin, 31 (4), 571-582, 2005.

Marcomini, A., Suter, G.W. II, Critto, A. Decision Support Systems for Risk Based Management of Contaminated Sites, Springer Verlag, New York, p. 436, 2009.

Margalef, R. The science and praxis of complexity. Ecosystem: diversity and connectivity as measurable components of their application. Pages 228-244 in Aida et al. (Ed), United Nation University, Tokyo, 1984.

Posthuma, L., Suter, G.W. II, Traas, T.P. Species Sensitivity Distribution in Ecotoxicology, Lewis publisher, Boca Raton, FL, USA, 2000.

Rutgers M., Den Besten P., Approach to legislation in a global context, B. The Netherlands perspective – soils and sediments. In: Environmental toxicity testing, Thompson K.C., Wadhia K., and Loibner A.P., eds., Blackwell Publishing CRC Press, Oxford. pp. 269-289. 2005.

Semenzin E., Carlon C., Critto A., Rutgers M., Marcomini A. Development of a site-specific Ecological Risk Assessment for contaminated sites: Part II. A multi-criteria based system for the selection of bioavailability assessment tools. Science of the Total Environment, 379, 34-45, 2007.

Shannon, C.E., Weaver, W. The mathematical theory of communication, University of Illinois Press, Urbana, IL, USA, 1949.

Shim, J.P., Warkentin, M., Courtney, J.F., Power, D.J., Shards, R., Carlsson, C., Past, present and future of decision support technology, Decision Support Systems, 33, 111-126, 2002.

Ten Brink, B.J.E., Hosper, S.H., Colijn, F. A quantitative method for description and assessment of ecosystems: The AMOEBA-approach, Marine Pollution Bull, 23, 265-270, 1991.

Von der Ohe, P.C., Prüß, A., Schäfer, R.B., Liess, M., De Deckere, E., Brack, W. Water quality indices across Europe – a comparison of the good ecological status of five river basin, Journal of Environmental Monitoring 9, 970-978, 2007

U.S. EPA, 1998. Guidelines for Ecological Risk Assessment- Final, EPA/630/R-95/002F, United States Environmental Protection Agency, Risk Assessment Forum, Washington, DC, USA.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

The implementation of EcoRA methodology in national-wide biomonitoring system of surface waters

J. Jarkovský[a,b], K. Kubošová[a,b], K. Brabec[a,b]

[a] *Research Centre for Environmental Chemistry and Ecotoxicology, Faculty of Science, Masaryk University, Kamenice 3, 634 00 Brno, Czech Republic (jarkovsky@iba.muni.cz)*

[b] *Institute of Biostatistics and Analyse, Masaryk University, Kamenice 3, 634 00 Brno, Czech Republic*

Abstract: Monitoring of water organisms communities has become a standard approach in surface water monitoring as well as a part of complex systems for assessing surface water quality. In European countries, the most commonly used organisms are water macroinvertebrates; developed in the United Kingdom, RIVPACS has been one of the first complex systems based on macroinvertebrates. Similar systems have been used in many countries worldwide and development of similar systems is also connected to the European Water Framework Directive. The European Water Framework Directive requires the good environmental state of the water bodies in the member states of the EU. The projects of the Ministry of Environment of the Czech Republic covers all aspects of this problem from data sampling to computer (informatics) solution. The general framework of the project respects the EcoRA methodology which provides the optimal solution for this type of environmental projects.

The EcoRA problem formulation is based on the requirements of the WFD ES which define time and spatial scale of the problem, priority stressors and key receptors for the monitoring. The exposure assessment is represented by concentration measurements of priority compounds for the evaluation of chemical status and sampling of biological communities for the evaluation of ecological status. The monitoring network covering both the spatial and seasonal aspects of surface waters was established for the measurements of stressors. The effect assessment is based on maximum allowed concentrations of chemicals in the environment for the assessment of chemical status. The assessment of ecological status requires the reference monitoring network which provides the expected natural values of multimetric biological indices and community composition with the similar meaning and application like maximum acceptable concentrations for evaluation of chemical status. For the evaluation risk based on chemical data the ratio of measured and maximum allowed concentration is adopted. The evaluation of ecological status is based on EQR ratios, i.e. the ratio of observed and expected ("natural") biological indices and community composition. The final risk evaluation integrates the results of both chemical and ecological status on the basis of "one out – all out principle". The evaluation of risk and its quantification or in terms of WFD the quality classes of status of surface waters are communicated through web visualization and reports. The results are applied for the identification of water bodies at risk and the planning of remedial measures for water ecosystems.

The EcoRA methodology is an important source of inspiration for the development of projects like this, i.e. systems for the evaluation of quality of the environment. It is able to identify and prioritize the necessary steps for the development of the system and transform them into strictly and clearly defined problem formulation, identification of data sources, risk identification and evaluation and finally the generalization, interpretation and communication of the results.

Acknowledgements: The presentation was supported by the Czech Ministry of Education, Youth and Sport (MSMT 0021622412 INCHEMBIOL).

Keywords: Water Framework Directive, benthic macroinvertebrates, biomonitoring, surface waters, predictive modelling, multimetric model, evaluation of ecological status

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Indicative value of benthic maroinvertebrates in implementation of WFD

J. Jarkovský[a,b], S. Zahrádková[c], K. Brabec[a,b], K. Kubošová[a,b], J. Bojková[c], P. Bartůšek[a,b], V. Syrovátka[c]

[a] *Research Centre for Environmental Chemistry and Ecotoxicology, Faculty of Science, Masaryk University, Kamenice 3, 634 00 Brno, Czech Republic (jarkovsky@iba.muni.cz)*

[b] *Institute of Biostatistics and Analyse, Masaryk University, Kamenice 3, 634 00 Brno, Czech Republic*

[c] *Department of Botany and Zoology, Faculty of Science, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic*

Abstract: The necessary component of any ecological risk assessment based on biomonitoring is the definition of susceptible organism reflecting changes in environmental conditions and disturbance of the environment.

In the evaluation of ecological status of surface waters required by the WFD the suitable monitoring network together with the correct selection of parameters including biological taxa determines the correctness and financial effectiveness of the monitoring. The selected biological endpoints have some prerequisites (common presence, consistent relationship to environmental condition, and susceptibility to stressors) which can be validated by means of analysis of indicative power of taxa and their ecological traits together with expert opinion.

The presented results are based on extensive databases of monitoring and biomonitoring data of surface waters in the Czech Republic since 1996 and answer the following questions: 1) What are valence characteristics of monitored taxa in undisturbed environment; 2) Ability of taxa to reflect stressors in the environment; 3) Validation of „species traits“characteristics which are utilized in the systems of evaluation of ecological status.

The final results of the project are numerical evaluation of indicative power of monitored taxa and identification of the most relevant taxa for routine monitoring, i.e. taxa with both consistent relationship to abiotic characteristics unchangeable by human activities and sensitive response to stressors in the environment. These results will be utilized in the implementation of the Water Framework Directive in the Czech Republic.

Acknowledgements: The project was supported by the Ministry of the Environment of the Czech Republic (project SP2e75008 VALENCE) and by project MSMT 0021622412 INCHEMBIOL.

Keywords: bioindicators, Water Framework Directive, benthic macroinvertebrates, biomonitoring, surface waters

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE : Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pilmmann, I.Holoubek, R. Legat (eds.)
Masaryk University, 2009

Risk mapping of technological risks (chemical, biological and aircraft) by means of GIS Risk-Register, the example of "Geneva Risk".

Alberto Susini[a] , Christian Hürzeler[b], Andreas Schönenberger[b], Alexandre Massot[c], Hartmut Fricke[d], Norbert Gronak[d], Corinne Tournier[a],

[a]Environmental compliance of companies, Canton of Geneva, Switzerland
(alberto.susini@etat.ge.ch)

[b]Basler&Hofmann, Consulting Engineers, Zurich, Switzerland
(andreas.schoenenberger@bhz.ch)
[c]Geneva information technology center (CTI), Canton of Geneva, Switzerland
[d]Gesellschaft für Luftverkehrsforschung GmbH (GfL), Berlin, Germany
(fricke@gfl-consult.de)

Abstract :The goals of the project were to obtain interactive maps for different kind of risks of industries submitted to risk regulations (chemical and biological risks) and to evaluate the interaction with other technological risks like aircraft damage. The results would then permit the setting up of priorities in monitoring installations and further enhance land-use practices and the preparedness of emergency response planning.

Keywords : Chemical risk assessment, classification and prioritization of chemical and biological hazards (IAEA methodology), cartographic representation of environmental safety and risk, aircraft accident damage, aircraft crash probability, aircraft crash consequences on hazardous facilities.

1. INTRODUCTION

The goals of the project were first to map different kind of risks (chemical and biological and then to evaluate how the risks for different types of industry, in particular those dealing with dangerous chemicals, may increase due to aircraft operations inside the area of the canton of Geneva. This evaluation would then permit the setting up of priorities in monitoring installations and further enhance land-use practices and the preparedness of emergency response planning. Special GIS representations were developed on the basis of these risk calculations values and special procedures will be put in place in order to inform the public.

The Geneva Risk project is an instrument jointly developed by Basler & Hofmann and the Geneva Information Technology Center (CTI) at the request of the Authorities of the Canton of Geneva to survey and to assess the risk of environmental potential pollution and potential human hazard caused by dangerous chemical substances and materials. In addition, Geneva Risk is a mean of implementing the Swiss Federal Ordinance on Protection against Major Accidents (Ordinance on Major Accidents, OMA) which deals with the handling of hazardous chemical substances and materials. Geneva Risk is also intended to be compatible with the European Community Seveso II Directive on the Control of Major Accidents

Hazards Involving Dangerous Substances, and also with the UNECE Convention on the transboundary effects of industrial accidents.

The project was further enhanced inside the Geneva-Risk system by the development of two new layers dealing with the legal compliance of the Ordinance on Contained Use of Organisms (CO) who regulates activities involving the contained use of genetically modified or pathogenic organisms in Switzerland, and with aircraft crash onto an industrial site. Risk assessments currently performed within the frame of the Swiss Federal Ordinance on Major Accidents (OMA) shall also take into account the consequences of highly unlikely events, including the extremely low probabilities of an aircraft crash onto an industrial site. Similar requirements are also tackled in the revised Seveso directive 96/82/EU.

Other existing data sources, such as the Geneva water information system, the Geneva social liabilities Database (provided by the Repertory of the Geneva Companies Competence Center (REG)) were improved in order to accomplish the objectives of Geneva Risk. This project was supported by the Geneva Geomatic Competence Center (who provided a wide range of geographical data with their database (SITG).

Our goals were to get a comparison of the different kinds of industrial risks in the canton, in order to set up priorities in monitoring installations, and to enhance further the preparedness of emergency response and land-use practices. Special GIS representations will be used in order to inform the public.

Risk assessments currently performed within the frame of the Swiss Federal Ordinance on Major Accidents (OMA) shall also take into account the consequences of highly unlikely events, including the extremely low probabilities of an aircraft crash onto an industrial site. Similar requirements are also tackled in the revised Seveso directive 96/82/EU, whereas until recently, such probability data was so far only determined for Switzerland in a single case, a joint study (EWI 1993) done for the Federal Office on Civil Aviation (BAZL) and the Federal Office of Military Aviation (BAMF). Hence, only generic aircraft crash probabilities were calculated for the entire area over Switzerland in this study. These values were expressed as the probability per year and square meter for both accidents during the en-route and the approach flight phase. Further, only four aircraft categories were considered for modelling. This is seen as not sufficient with regard to worldwide aviation accident data of the last decades showing different aircraft crash statistics, this mainly caused by the fact that aircraft are not randomly distributed in the airspace but follow standardised procedures especially during taking off and landing. A serious attempt to systematize the data processing of aircraft accident into hazardous facilities was made in 1996 by the US Department of Energy (DOE 1996). When evaluating the adoption of their methodology to the situation around European Airports, it became clear, that various airport conditions and operational modes could not give a sufficiently reliable determination of aircraft crash frequencies. Potential aircraft accident risks in the vicinity of hazardous facilities situated near the Geneva International Airport would consequently not be sufficiently taken into account in the local risk management and land use practices.

The cartographic representation of hazards due to aircraft supplements the wide range of tools and projects forming part of Geneva Risk, a project initially dealing only with potential risks and damages caused by dangerous chemical substances (Susini et al. 2004; Hansen et al. 2005) and later completed with biological risks (Susini et all. 2006 and 2007a). GIS methods make it possible to combine and sum up the risks derived from individual storage or process events into a cumulative risk for the plant. GIS also allows to visualize the risk levels and to partition them to specific areas. Furthermore, GIS facilitates the interpretation of data and presentation of the final numerical results. The accumulated risk layers of aircraft accident with major hazard plants lead to a prioritization of plants susceptible to involve significant risks to the population in the case of aircraft accidents. All of this was made possible by automating the interaction between an Oracle database, and various different GIS - layers.

Risk visualization using GIS is used to identify risk clusters and to facilitate control and law enforcement, as well as for the prevention and mitigation of the effects of possible future accidents .

2. METHODS

2.1 METHODOLOGICAL BACKGROUND

The Cantonal ICT infrastructure is made up of different units which interact strongly. The communication infrastructure is managed by the Geneva Technology Information Center (CTI), the Geneva liabilities data are given by the Geneva Companies Competence Centre (REG; http://reg.ge.ch). linked with the Federal Companies Centre (BUR-REE; http://www.admin.bfs.ch/), a wide range of standardized cartographic data are taken from the cantonal geographic database (SITG; http://www.sitg.ch) provided by the Geomatic Competence Centre (GCC). Finally, the information system of the environmental affairs section of the Geneva Labor Inspectorate follows up the technical installations of companies referenced in the Cantonal Risk Register (CRR). This CRR is used to manage the risk calculation data of industrial facilities with chemical and biological risks to allow for cartographic visualization, and also for monitoring and administrative follow-up according to legal compliance of all major hazard facilities situated in the Geneva canton.

Numerical models were defined and applied to the release and dispersion calculations for a range of sources of risk for the population and the environment, in particular the stationary facilities (storage and production of chemicals and contained use of pathogenic organisms) and the transport infrastructure (motorways, main roads, railways and pipelines). The Geneva Risk project, deals mainly with the risk on industrial installations which fall under the Ordinance on Major Accidents (OMA) (Susini et al. 2004, Hansen et al. 2005, Susini et al 2006, Susini et al. 2007a, Susini et al. 2007b). Further developments were made with the conception of new calculations allowing to estimate the aircraft crash consequences on hazardous facilities (Susini and al. 2008). These first developments made with a joint venture with the canton of Zurich (Chemrisk project) were then further enhanced by other Swiss cantons (Aargau, Turgau, Lucern and Basel-Stadt) joining this initial partnership and a special geomatic interface called RCat was developed by Basler&Hofmann. The RCat ArcGIS interface makes it possible to import and perform calculations necessary for cartographic risk mapping more efficiently.

2.2 IT ASPECTS FOR CARTOGRAPHIC PROJECTION OF RISK CALCULATIONS

To handle the data forming part of the digital infrastructure of the Geneva canton, the use of relational databases for the data administration (industrial facilities and their associated hazard potentials and risks) is important in order to allow for systematic calculations.

It was decided to use an internationally recognized standard, based upon three hierarchical descriptive levels (Organization, Site and Entity) as defined by the European Union EMAS eco-audit directive (EC No 761/2001), together with a decision form specified by them, on the 7 September 2001, for the implementation of the Regulation (EC No 761/2001). The precise meanings of the descriptive levels is as follows :

"Organisation" shall mean a company, corporation, firm, enterprise, authority or institution, or part or combination thereof, whether incorporated or not, public or private, which has its own functions and administrations.

"Site" shall mean all land at a distinct geographic location under the management control of an Organisation involving activities, products and services. This includes all infrastructure, equipment and materials.

"Entity" shall mean a part of a site or subdivision seeking to register under one registration number. This includes for example large amounts of chemicals situated in storage and production units located inside an industrial site.

The descriptive level of the site has a key number attribute, initially developed for statistical and fiscal duties that were also used to interface with other available and quality-controlled data, such as the federal building register and the geographic values required to create GIS cartographic representation. The entity descriptive level, , was used for the calculation data. This entity represents, inside an industrial site, a storage or processing part containing significant amounts of chemical substances. In the database this level is geographically referenced by cartesian (x,y) coordinates. Technically speaking, for the Geneva Risk project a Spatial Database Engine (SDE), developed by the Environmental System Research Institute (ESRI), was established to serve the varied needs of different administrative units. The SDE server provides a variety of raster-based information, such as satellite images and maps with different scales, and vector-based data, such as river and highway specifications. Security aspects and personal data protection rules were duly taken into account. There are separate sets of databases accessible to the public, to internal administrative units and to cantonal agencies. The calculations made within the database precisely determine the accidental damage extent of an entity in terms of percentage of affected people with mortality probability indicators (10% and 100%). The calculations and accidental frequencies specific to the entities then form part of several files which are then processed in the RCat interface in order to create visual models in the form of maps. The RCat interface processes the data, together with the Geneva cartographic server (SITG) demographic data in the GIS), in order to calculate the exact extent of the consequences (death,) to people, this result is then multiplied by the accident probability in order to obtain the corresponding risk calculation, In case of an industrial accidental event, according to the specific accident scenario, the probabilities are given by the CRR; in case of an aircraft accident, the probability used is stored inside the RCat interface. The results of these calculations is then exported by RCat in order to be integrated into the the overall data of the CRR. The ArcGIS software interface was used to implement the risk assessment output for visualization purposes. The different file exchanges between the CRR Oracle database and the RCat ArcGIS system are depicted in figure 1.

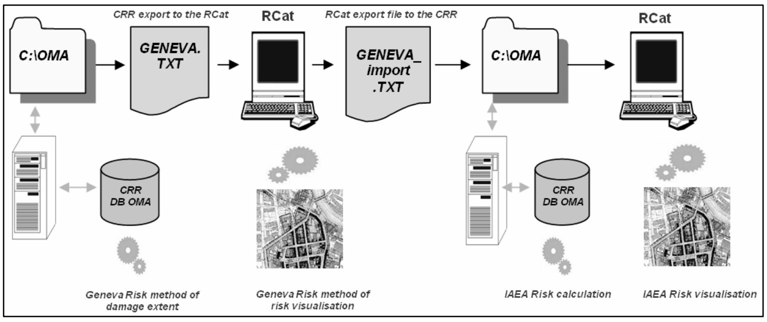


Figure 1: file exchange between the CRR and the RCat for calculations and map visualizations.

There is no method available to quantify the risk of handling pathogenic micro-organisms in clinical, research, and environmental laboratories, production plants etc.. Based on the IAEA-approach to rapidly assess and prioritize industrial risks, a simple risk model to quantify the risk of pathogens to lab workers and their contact persons outside the facility was developed for different types of activities like diagnostic, food diagnostic, research laboratories, and production installations within the framework of confined use legislation.

The calculation for the biological risks were made possible by the intent to automate biologic risk management data exchange between the Swiss federal information system (ECOGEN) and the Geneva cantonal information system on the industrial environment. Using adapted sorting criteria, all the selected projects on the contained use of pathogens were submitted to risk calculations as part of the application of the Geneva risk project called the

Cantonal Risk Register incorporating exports for GIS risk representation (Susini et al. 2006, Susini et al. 2007a).

The basic hypothesis is that in the case of an aircraft accident on a specific entity containing chemicals stored or processed inside that site, the full quantity of the substance is released into the environment. This is clearly a worst-case approach with regard to typical industrial accidents in which specific scenarios envisage only parts of the chemicals being released. These calculations were automatically performed for a large number of plants referenced inside the cantonal risk register (Susini et al. 2008).

3. RESULTS AND VISUALIZATION

Without visualization the results of this project would be hard to use, communicate or explain. So visualization was one of the key goals. The following figures show how risk assessment results may be illustrated (fig. 2, 3 and 4,). Figure 2 shows how potential damage of plants can be located, figure 3 shows how individual risk maps can give an additional information for land use planners for example, and at least an example of biosafety relevant image for decision making process and risk comparison, since the data base for the enterprises includes a feature that allows to plot each enterprise within a diagram of an FN-curve (Frequency of N or more fatalities, as function of N) (fig. 4).

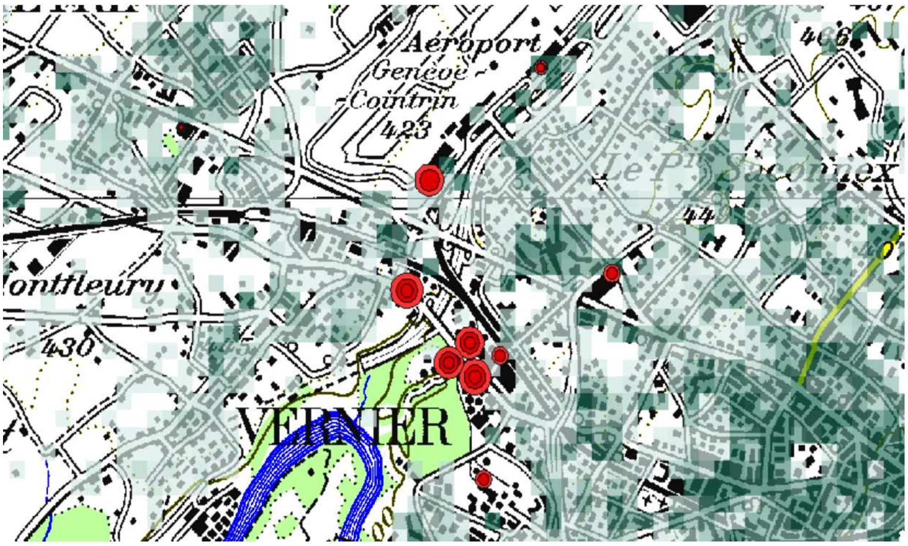


Figure 2 : Plants near the Vernier area with damage potential (fictive results).

Figure 3 : Plants near the Vernier area with individual risk contours (fictive results).

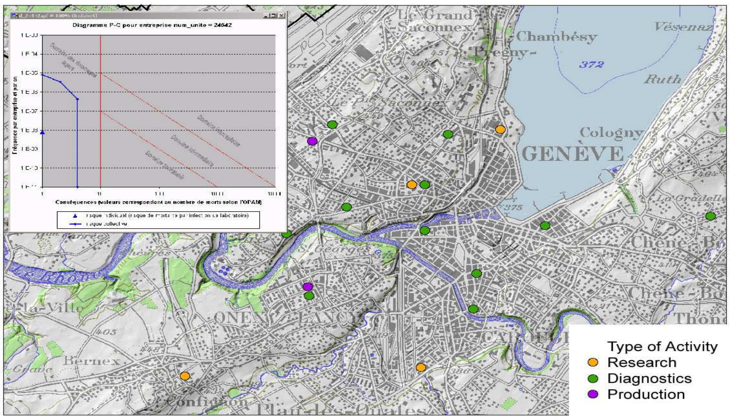


Figure 4 : Cartographic representation of biohazards in containment facilities (laboratories, industrial plants).

The results made of the calculations performed with the aircraft data, made it possible to identify those sites collections for which air traffic operations do have a negative impact on the accidental risk. If produced only as tables of numerical values, the results of this project would be hard to use, communicate or explain. A graphical presentation was thus one of the key goals. The following two figures show how risk assessment results may be clearly presented. Figure 5 shows the potential risks to plants induced by aircraft accident based on their location.. Figure 6 shows the specific risk contributions of air traffic operation hazard related to Geneva Airport and those due to normal accident scenarios. As might be expected, the results show several cases where there is a significant increase in risks for plants situated located closely a given departure or arrival flight route from or to Geneva Airport. The results showed that our approach allowed us to prioritize correctly the concerned major hazard facilities around the Geneva airport. These results confirmed the correct sensitivity of the chosen methodology who gives a predominant influence of the aircraft accident probability.

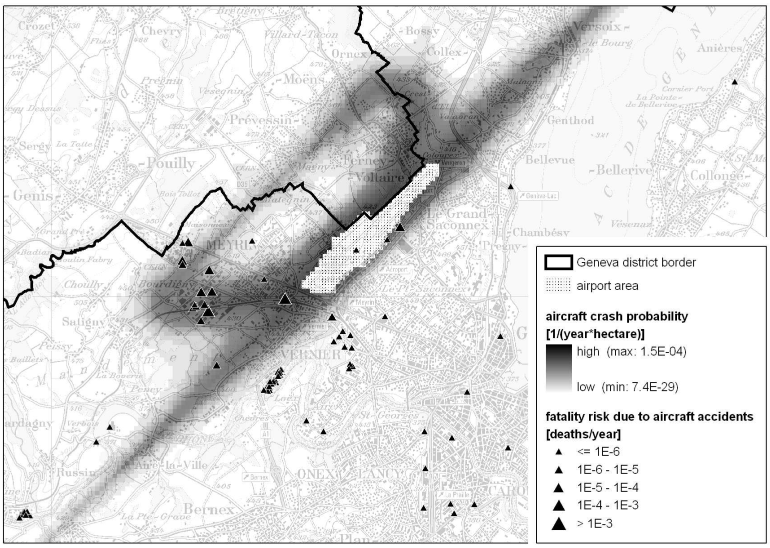


Figure 5: Potential risk of plants due to aircraft accidents (fictive results)

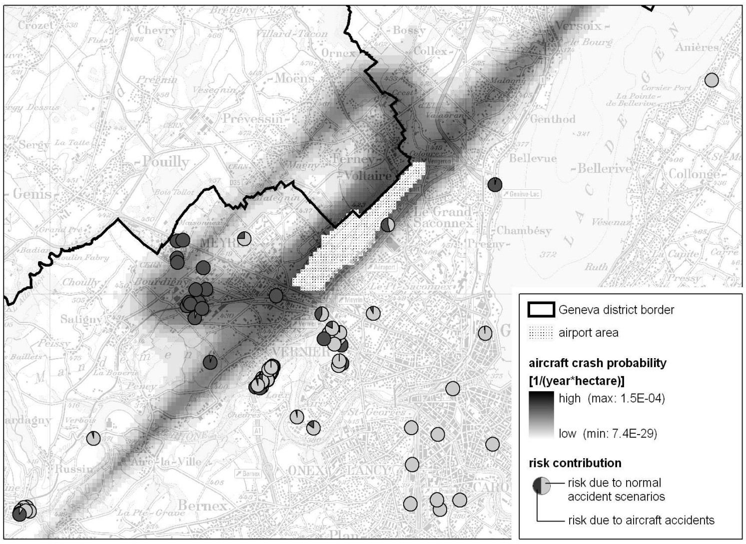


Figure 6: Risk contribution (risk due to aircraft accidents in comparison to risk due to normal accident scenarios) (fictive results)

4. CONCLUSION

The cartographic representation of hazards due to aircraft supplements the wide range of tools and projects forming part of Geneva Risk, a project initially dealing only with potential risks and damages caused by dangerous chemical substances (Susini and al. 2004; Hansen and al. 2005) and later completed with biological risks (Susini and al. 2006 and 2007a). GIS methods make it possible to combine and sum up the risks derived from individual storage or process events into a cumulative risk for the plant. GIS also allows to visualize the risk levels and to partition them to specific areas. Furthermore, GIS facilitates the interpretation of data and presentation of the final numerical results. The accumulated risk layers of aircraft accident with major hazard plants lead to a prioritization of plants susceptible to involve significant risks to the population in the case of aircraft accidents. All of this was made possible by automating the interaction between an Oracle database, and various different GIS - layers. The lists of dangerous chemicals and organisms with their properties will be part of the scientific support in the event of spills involving hazardous materials.

Risk visualization using GIS is used to identify risk clusters and to facilitate control and law enforcement, as well as for the prevention and mitigation of the effects of possible future accidents .

The main beneficial results of the project can be summarized as :

- The projections could be shared with other governmental units in the Canton of Geneva, in particular land management involving permits regarding storage of toxic gas inside sensible major hazard plants.
- Inspections of industrial facilities were prioritized based on the Geneva Risk decision-support tools.
- Safety improvement measurements were evaluated
- A map for information of the public is now available.

- The results calculated and visualized by means of GIS are judged to be an important instrument of risk communication between stakeholders, authorities, public and the airport

REFERENCES

DOE : Accident analysis for aircraft crash into hazardous facilities. DOE Standard. DOE-STD-3014-96, 1996

EMAS 2001/681/EC: Commission Decision of 7 September 2001 on guidance for the implementation of Regulation (EC) No 761/2001 of the European Parliament and of the Council allowing voluntary participation by organisations in a Community eco-management and audit scheme (EMAS) 2001 http://europa.eu.int/comm/environment/emas/documents/legislative_en.htm

EWI: *Studienauftrag : die Auswirkung der Luftfahrt auf die Umwelt, Technischer Bericht F, Teilbereich Absturzrisiken.* Elektrowatt Ingenieure im Auftrag von Bundesamt für Zivilluftfahrt (BAZL), Bundesamt für Militärflugplätze (BAMF), 1993.

IAEA : Manual for the classification and prioritization of risks due to major accidents in process and related industries, International Atomic Energy Agency (IAEA), Vienna Nov, 1996, IAEA-TECDOC-727 (Rev. 1), 1996.

Hansen, J., Funk., W, Susini, A., Gmünder, F., Hürzeler, C.,Knechtle, N., Schönenberger, *A Management of major chemical hazards by means of a geographical information system.* 14th SRA-Europe-Annual Meeting, Como (IT), 12-14 Sept. 2005.

Susini, A., Knechtle, N., Perrin, M., Hürzeler, C., Gmünder, F., Labarthe, R., Minier, Ph, Gay, B. : Sharing Environmental, Spatial and Technical Data for Risk Assessment within the Geneva Region, « Geneva Risk », in Proceedings of the EnviroInfo 2004 Conference, pp. 363-370, 2004.

Susini, A., Gmünder, F., Maliki, G., Tournier, C. : *Highly Integrated Information Systems for Automated Biological Risk Assessment of the Contained Use of Pathogenic Organisms within the Geneva Region*, in Proceedings of the EnviroInfo 2006 Conference, pp. 417-424., 2006.

Susini, A., Gmünder, F., Barres, L., Tournier, C. (2007a): Complexity of modeling biological risks with map production of *Contained Use of Pathogenic Organisms in facilities of the Geneva Region*, in Proceedings of the ISGI 2007 Symposium on generalization of information, Codata-Germany, pp. 158-177, 2007.

Susini, A., Hansen J., Gmünder, F., Schönenberger, A., Hürzeler, Ch., Minier, Ph., (2007b): *Cartography of automated environmental technological risks by means of GIS Risk-Register, the examples of Geneva, "Geneva Risk" and Zurich "ChemRisk*", in Proceedings of the EnviroInfo 2007 Conference, pp. 655-662, 2007.

Susini, A., Hürzeler, Chl., Schönenberger, A., Massot, A., Fricke, H., Gronak, N., Thiel, Ch., Fiedler, D. (2008) Assessment of aircraft accident probability on industrial facilities by means of GIS Risk - Register, the examples of Geneva "Geneva Risk" in Proceedings of the EnviroInfo 2008 Conference, pp. 466-475, 2008.

Workshop 4

eEnvironment Terminology

Organized by **Thomas Bandholtz**, Joachim Fock, Rudolf Legat,
Michal Nagy and Paolo Plini

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T.Bandholtz (eds.)
Masaryk University, 2009

Meaningful Linkage and Navigation in the Shared Information Space

T. Bandholtz

innoQ Deutschland GmbH, Halskestr. 17, D-40880 Ratingen, Germany (thomas.bandholtz@innoq.com)

Abstract: SISE intends to use "semantics, thesauri, ontology services and standardisation" in order to support "flexible chaining of distributed environmental services". How can this be achieved? Formalised Semantics provide powerful means for any kind of data and service integration . The environmental community looks back to a large experience with vocabularies and metadata-driven access to information. Learning from the latest integration patterns implemented by the Semantic Web community, we can build a meaningful linkage and navigation in the upcoming Shared Information Space.

Keywords: *Data; Metadata; Navigation; Ontology; Semantic Web; Thesaurus.*

1. INTRODUCTION

Looking back to some thirty years of semantic experience in the environmental community, summing up the hard and stony way trying to overcome the gap between data and metadata, vocabulary and indexed information, one might come to the conclusion that we forgot only one thing: how to keep it simple.

2. EVOLUTION OF METADATA APPLICATION PATTERNS

Environmental information is a comparatively young discipline. It came up in the 1970s as a melting pot for highly engaged stakeholders with an exceptionally diverse background in agriculture, biology, economy, energy, engineering, ethics, geology, chemistry, management, medicine, meteorology, oceanography, regional planning, zoology, and more. All these actors got together by a serious concern about the sustainability of the natural environment: the common target was environmental protection, not primarily information or science. Nevertheless, all these actors had to learn to talk to each other first of all. There was no shared language among them beyond the language of complain, protest and campaign in the early years.

2.1 Librarian Subject Catalogues and Early Thesauri

There had never been any tradition (or even a notion) of something like "*environment*" in science or in any governmental portfolio before. When one of the first governmental departments was founded in 1974 (the Federal Environment Agency) in Germany, not surprisingly they established an interdisciplinary library which today holds "some 300.000 books, approximately 1.000 periodicals and press releases as well as microfilms, 170.000 American research reports on microfiches and more" (UBA 2008).

One of the main challenges was to maintain an appropriate subject catalogue covering all the contributing disciplines. This led to the development of an "Environmental Thesaurus" (UMTHES®) which today holds some 10,000 descriptors and 25,000 non-descriptors (bi-lingual German/English). Similar processes took place in many other countries resulting in vocabularies such as "*Thesaurus Italiano per l'Ambiente* (TIA)", "*Thesaurus de Medio*

Ambiente" in Spain, "*Lexique environnement - Planète*" in France, as well as multilingual approaches such as the "*Multilingual Environment Thesaurus* (MET)" from the Netherlands or the "*EnVoc* Thesaurus", of UNEP Infoterra.

Obviously, this “melting pot” of contributing sciences generated an exceptional interest in interdisciplinary vocabularies which remained unbroken till today.

2.2 Digital Data Catalogues

The early nineties brought a wide spread use of computers and huge collections of heterogeneous observation data. Metadata made the move to *digital* representation. and to the description and categorisation of data sets (the “data catalogues”). This was the era of metadata *profiles* in the environmental domain, some of which had been established and harmonised (at least on a national level) even before the DCMI was founded in 1994. The main obstacle was that any direct access to the data behind the description was still missing. The contributors were not yet wired, and even the metadata had to be distributed on disk (if not via some few expensive leased circuits). Usually the access point was a postal address. Such collections were valuable inventories of collected data, but they suffered from slow update and distribution procedures and sometimes intricate negotiations preceding (or hindering) any final data access.

In these years also the thesauri became digitized, but not yet networked. All through this development the vocabularies were hosted by the information systems as kind of a built-in feature. Each system used its own “home” terminology (even if some used terminologies of the same origin) and the “keywords” were simply referenced by name. These names were forced to be unique within any “home” terminology, and the origin of this terminology remained implicit.

By this time also the European Environment Agency (EEA) was established in 1990 by the European Union, and became operational in 1994. This led to some international collaboration and harmonization quite soon. In 1996 the EEA initiated the GEneral Multilingual Environmental Thesaurus (GEMET) “aimed to define a common general language, a core of general terminology for the environment” which has been compiled by merging selected terms from eight different Thesauri from Germany, Italy, Netherlands, Spain, France, and European and international organisations. GEMET has been maintained by the EEA only until 2001 but is still available and widely used with ca. 5300 terms, each of them with (symmetrical) translations in 21 languages.

In 1998 the European Topic Centre on Catalogue of Data Sources was established in order to integrate the early approaches on a European level. After some initial success it has been finalized for not so obvious reasons in 2001. The closing Workshop in Thun (CH) became the starting point of the DCMI Environmental Community in order to provide some continuity to the participants.

2.3 Web Indexing

With the dissemination of the Internet in the late 1990s the focus moved to Web Indexing, and this also led to some first vocabularies becoming available over the network for everyone.

By the same time, the Aarhus Convention on “Access to Information, Public Participation in Decision-making and Access to Justice in Environmental Matters” (Aarhus, 1998) was signed in 1998, which introduced a continuous liberalisation of public data access and finally to an obligation to active disclosure of information.

In this way the technical opportunities and the administrative obligations joined forces to generate a first wave of public Internet representations of the environmental authorities. When the German Environmental Information Network (Bös, 2000) went into production with a full text search engine (ht://dig) in 2000, it crawled more than 80,000 Web pages of 50 information providers (today’s follow-up “PortalU” is dealing with over a million, of course). Suddenly the information itself was no longer hidden in libraries or on disks, but became directly accessible (and searchable!) in a more and more common infrastructure.

The power of full text Web indexing suddenly became a critical challenge for the metadata and the thesaurus community. The question to be answered was: “Can we do better than Google?” GEIN maintained two indices, one of them thesaurus-based, so the public and professional users could compare the results immediately. The professionals clearly appreciated a controlled vocabulary, though the resulting information was not in each case more appropriate. Most of the public users however did not even understand why they should navigate some terminology before searching the information itself.

Regarding a metadata profile, GEIN used a very simple set of title, description and URL, along with subject headings from the thesaurus, place names, and temporal extent of the information. Even this was considered dispensable by many participants. An attempt to place such metadata within HTML meta tags was dropped very soon, as none of the information providers really cared about it. GEIN, and later SNS, relied on their own analysis of the original document sources in order to generate and store such annotations by themselves.

In 2001 the Federal Environment Agency started the Semantic Network Service (SNS) project. (Rüther 2008). The intention was to outsource and enhance the handling of vocabularies from GEIN into a specialized Web service that would serve not only one information system but the environmental community in whole. SNS provides a complex, bilingual vocabulary organised as a Topic Map (ISO 13250) and several semantic methods that can be accessed by Web Services:

- findTopics: Search for topics by names and topics types..
- getPSI: Provides a reference of each topic with characteristics and associations (PSI=”Published Subject Indicator” of the Topic Map standard).
- autoClassify: supports automatic indexing of plain text, (X)HTML, and PDF documents, resulting in a ranked list of significant topics covered by the document.
- getSimilarTerms: returns a “cloud” of synonyms for a given search term as a suggestion for a more successful full text search & retrieval.
- getHierarchy: a composition (replacing multiple getPSI requests) in order to retrieve a narrower/broader hierarchy from a given Thesaurus descriptor
- findEvents: simply enhances findTopics with temporal search conditions.
- anniversary: finds events in the environmental chronology that happened x years ago by a reference date.

These services are supporting currently nine environmental / spatial information portals, among them PortalU and GeoPortal.Bund, which must be seen as a proof of concept at least. SNS will also be used to provide keyword values for German contributions to INSPIRE, starting in summer 2009.

However, the benefits of a *controlled* vocabulary remained a contentious issue till today. The most popular use of vocabularies today is looking up synonym terms in the context of full text search.

2.4 Geographic Service Annotation

The next move of environmental metadata - which is currently ongoing - was focused on geospatial references, and it introduces the style of service annotation.

As most of the environmental data collections are in some way geo-referenced, the environmental authorities have always observed and contributed to any kind of geographic location service, starting with the early *Global Information Locator Service*, GILS, and the *Global Environmental Information Locator Service*, GELOS, (Christian) moving on to the

Global Monitoring for Environment and Security (GMES) or the *Global Earth Observation System of Systems* (GEOSS) today.

The most effective of such geospatial approaches in the environmental domain in Europe currently is the *Infrastructure for Spatial Information in the European Community* (INSPIRE). Compared to Web indexing, INSPIRE can be regarded as the return of the metadata profile. INSPIRE includes *OGC Catalog Services* (OGC CAT) which again are based on ISO 19115 *Geographic information -- Metadata*. ISO 19115 defines some 20 mandatory core attributes plus over 400 optional properties, and it allows for *community profiles* that may even add their specific properties not contained in the standard. In this it is rather similar to the DCMI model, however obviously there has not been any consultation or alignment between the two bodies. The only contribution to be found in this context is an ex-post *Mapping between Dublin Core and IS0 19115* (CWA 14857), as a result of the *Meta-Data (Dublin Core) Workshop* by the European Committee for Standardization, CEN, in 2003. This mapping tries to “improve the discovery of geographical information in cross-domain searches”, including a “style sheet for automatic translation mapping between ISO-XML and DC-RDF”.

While OGC CAT is a service by itself, its role is to provide metadata describing two different kinds of OGC services, namely the Web Feature Service (WFS) and the Web Mapping Service (WMS), each of them carrying some kind of data collection. That is why this pattern may be called “service annotation”. The target of the metadata is no longer some “hidden” data, and it is not a Web page, but it is a service that directly gives access to the described data. Exactly this is something that the early data catalogues had aspired after, but could not realize due to a lack of standardization coverage.

If INSPIRE/OGC CAT may be seen as the return of metadata profiles in the environmental domain, it also may provide considerable stimulation on controlled vocabularies. The INSPIRE Metadata Implementing Rule (INSPIRE, 2008) stipulates the use of multiple “Keyword” attributes along with an indication of the “Originating Controlled Vocabulary” for each of them.

Specific for the environmental domain, GEMET has been chosen as the preferred controlled vocabulary within INSPIRE which certainly will lead to a reanimation of liable maintenance in the near future, after several years of homelessness.

2.5 Shared Information System and Single Information Space

While these implications of INSPIRE are gaining awareness by the stakeholders, the next generation of environmental information is already waiting in the wings. Evidently not every piece of environmental information is geo-referenced, and a spatial-oriented discovery and presentation approach is not convenient for all purposes, even if “tagged” by keywords.

There has always been a large amount of location-independent information about topics like chemicals (substances), engines, management, production procedures, human health, climate change, consumption patterns, etc... Such non-geo-referenced information is getting more and more important delivering a general understanding of environmental aspects and supporting a new approach of integrated policies.

The “Communication on establishing a *Shared Environmental Information System* for Europe to improve and streamline the European system for collecting, analysing and reporting environmental information”, SEIS, has been adopted by the European Commission in early 2008 (European Commission, 2008). SEIS draws a vision “in which data would be a shared and accessible resource”, integrating the INSPIRE approach as well as many others. In SEIS “the current, mostly centralised systems for reporting are progressively replaced by systems based on access, sharing and interoperability.” From a technical point of view, the most crucial issues of this strategy are:

- “information should be managed as close as possible to its source”, and
- “information is provided once and shared with others for many purposes”.

So far SEIS does not reveal much about *how* "access, sharing and interoperability" shall be enabled, and how discovery of distributed systems shall be managed. SEIS seems to be oriented to some necessary "administrative streamlining" in the current phase.

Complementing SEIS, the ICT for Sustainable Growth Unit of the Commission's Information Society and Media Directorate-General is promoting a "*Single Information Space in Europe* for the Environment" (SISE) as a future research or innovation area within FP7. The current prioritization of topics in future calls is headed by "*Flexible chaining of distributed environmental services*" and contains several issues that are impacted by vocabulary management, such as:

- Methods and protocols for service discovery and chaining
- Semantics, thesauri, ontology services and standardisation
- Web 2.0 collaborative community and semantically enhanced services.
- Tools for interactive usable and useful contextualised user interfaces.

The specific role of metadata and vocabulary in this vision is not yet defined, but there is at least one dedicated participant within the SEISnet Community: The Semantic Interoperability Centre Europe (SEMIC.EU) provides a repository and a browser for shared "interoperability assets", namely XML schemas and Ontologies.

3. WHO CONTROLS CONTROLLED VOCABULARIES?

Among the various questions that arise from such metadata application styles as discussed above, we will focus on the challenge of vocabulary management, or, as some may call it: *ontology governance* in the following.

3.1 Controlled Vocabularies within INSPIRE / OGC

The INSPIRE/OGC CAT approach adopts the "Keyword" metadata element from ISO 19115. This allows for multiple keywords, each of them along with an optional "citation" of its "originating controlled vocabulary". ISO does not contain any restriction about the selection of such vocabularies. While it may be up to the creators of a DCMI application profile to specialize and extend the list of vocabulary encoding schemes, in the ISO 19915 approach the decision may be left to the provider of each metadata instance.

The INSPIRE Metadata Draft Implementing Rules (INSPIRE, 2008) state more closely that "*at least one keyword shall be provided […] from the general environmental multi-lingual thesaurus (GEMET)*", and "*If the keyword value originates from a Controlled Vocabulary (Thesaurus, Ontology, ...), for example GEMET, the citation of the originating Controlled Vocabulary shall be provided.*". Keywords from other sources are "also encouraged". The future will show how this gets handled.

GEMET itself can be accessed on the EIONET Web site[1] provided by the European Environment Agency, but there is neither a clear statement about ownership nor a dedicated editorial team since 2001 (see below).

Still the INSPIRE/OGC CAT data type of "keyword" is a simple string - it does neither support URL identifiers nor labels in different languages. The citation of the "originating controlled vocabulary" is of a complex type holdings postal addresses and several properties among which a URL may occur, but there is no way to express how to access a vocabulary service in order to explain what the keyword means and how it maps to different keywords from other vocabularies.

The original OGC vocabulary style follows the URN scheme as defined by IETF RFC 3406 *Namespace Definition Mechanisms*. URNs are another form of URIs, following an abstract

---

[1] http://www.eionet.europa.eu/gemet

naming scheme *not* bound to the HTTP protocol like URL, and thus not directly resolvable. OGC maintains a *Naming Authority*, OGCNA, which is still under development. URN from various sources are widely in use throughout the OGC service data, however this is neither well documented nor without controversy.

3.2 Controlled Vocabularies in the Environmental Community

Chapter 2 gave a broad introduction into the attention towards controlled vocabularies in this domain. This needs to be completed by a characterisation of the currently ongoing approach of vocabulary management. This approach may be summarized: Just like the environmental community is used to recognize the diversity of *species* as a "protected asset" it is likely to recognize the diversity of vocabularies in a similar way.

In 1996, the above mentioned early "librarian" Thesauri joined forces to publish the integrated GEneral Multilingual Environmental Thesaurus (GEMET), and there had been an international editorial team for some years, but it came to an end in 2001. Since then, there have been several new translations following the EU extension, but there has not been any extension of the vocabulary itself. For example, GEMET does not know "emission trading".

In 2004, the ECOInformatics Initiative launched a global "Thesaurus and Terminology" area of activity, called "Ecoterm", in order to facilitate the sharing and reuse of environmental information by "some level of semantic interoperability across these terminologies." (EcoInformatics). A survey counted more than 40 much larger vocabularies, many of them with an interdisciplinary and/or multilingual approach – but merely without any cross-references among each other so far.

Ecoterm held an annual meeting since then, trying to push four areas of activity with some first results.

(1) "*Implement Web Services connections to their terminology system that allow the EI Web Service to query the terminology in real-time*": In 2004 only the Semantic Network Service (SNS) already mentioned provided such Web Services. In the following years some others followed (among them GEMET), but a common pattern is still missing.

(2) "*Provide full compliance with the Terminology Web Services specification*": While several Ecoterm participants started to implement Semantic Web representations (SKOS and OWL) of their vocabulary, there was no clearance about service operations und only little awareness of RESTful http access.

(3) "*The resource will be described on the EI Terminology page with a link to the terminology, if it is publicly available*": Apart from publishing the results of the initial survey in 2004, there has been no kind of such a registry development till today.

(4) "*Develop a unique web service of interest to the group, for example, a web service that supports language translation or presentation of chemical structures …*": Except from the fact that many of the vocabularies had been multilingual by themselves right from the start no such service has been communicated to the group. However, there has been a forward-looking contribution by the Ontology Alignment Evaluation Initiative (OAEI 2007) who aligned three SKOS formatted thesauri (GEMET, AGROVOC, NAL) using relations from the SKOS mapping vocabulary.

Nevertheless, this challenge has not been solved yet although a lot of work in the right direction has been done. Handling keywords within INSPIRE and, for the future, providing semantic orientation and interoperability within SISE/SISE will stimulate these activities by providing a more practical application area with a more powerful obligation.

As already mentioned, INSPIRE has selected GEMET as the preferred vocabulary. Since then, there has been a first rudimentary approach to integrate GEMET with the INSPIRE Spatial Data Themes, but this has been withdrawn from the GEMET Web page, and there is no statement about the future editorial approach.

4. TOWARDS A SEMANTIC WEB OF ENVIRONMENTAL DATA

In his initial architectural plan for the Semantic Web, Sir Tim Berners-Lee drafted "*a web of data, in some ways like a global database*." (Berners-Lee, 1998). One of the latest W3C publications in this area, "Cool URIs for the Semantic Web", comes even closer to the notion of the environmental community: "*The Semantic Web is envisioned as a decentralised world-wide information space for sharing machine-readable data with a minimum of integration costs.*" (W3C, 2008). Sometimes it looks like this target has been pushed to the back seat by debates about formal languages, philosophy, and artificial intelligence topics such as interfering and reasoning. Vocabularies however play a crucial role in this information space, clarifying the meaning of data elements as well as providing "code lists" or "value sets" for the content encoding.

However, the communities are still missing a more strategic application pattern which links such vocabularies to digital libraries and to masses of machine-readable data. How could this become kind of a "global database", or "information space", in the end? A dedicated TaskForce of W3C (LinkingOpenData) has already managed to interlink about "two billion" items from more 30 different sources, based on four simple rules:

(1) Use URIs as names for things

(2) Use HTTP URIs so that people can look up those names.

(3) When someone looks up a URI, provide useful information.

(4) Include links to other URIs, so that they can discover more things.

Most of these principles are well known from the "Internet" which allows seamless navigation across distributed information. The main difference is denoted by using the term "thing" instead of "web page". From the perspective of vocabulary governance, such "things" are essentially concepts.

When someone looks up a concept URI, they should find a short description of the meaning, and they should find links to related data sets and reports.

Likewise the data sets (and any kind of report metadata) should use such concept URIs inside to link back to the vocabulary.

Whatever the user's starting point has been, they can navigate through such networked information following links between concepts (what is it about?) and related data (what has been measured?), exploring the context (places, methods), timelines, drivers, pressures, impacts, and agents.

This architecture leads to a Web of environmental data which may overcome the restrictions of fixed reports and "content management" approaches. From this point of view, the environmental community will have to rethink the role of metadata and vocabularies, and it looks like SEIS/SISE could finally become the appropriate platform to establish this evolving process.

REFERENCES

AARHUS. Convention on Access to Information, Public Participation in Decision-making and Access to Justice in Environmental Matters. Aarhus 1998. http://www.unece.org/env/pp/

Berners-Lee, Tim. Semantic Web Road map. W3C Design Issues. 1998. http://www.w3.org/DesignIssues/Semantic.html

Bös, Richard, Maria Rüther and Thomas Bandholtz. The German Environmental Information Network (GEIN). Environmental Informatics Symposion, Bonn 2000. http://enviroinfo.isep.at/UI%20200/Bandholtz-2000-07-18.el.pdf

Christian, Eliot. The Global Information Locator Service. Preprints of the Metadiversity Conference Proceedings. Charleston, SC, 2002. http://www.nfais.org/publications/metadiversity_preprints15.htm

EcoInformatics. Cooperation between USEPA and EEA. http://ecoinfo.eionet.europa.eu/

European Commission, DG Environment. (2008). Communication on a Shared Environmental Information System for Europe. http://ec.europa.eu/prelex/detail_dossier_real.cfm?CL=en&DosId=196687

GEMET. (2004). GEneral Multilingual Environmental Thesaurus. http://www.eionet.europa.eu/gemet/about

GEOSS. Global Earth Observation System of Systems (GEOSS) http://www.earthobservations.org/

GILS. Global Information Locator Service (GILS) - making it easier to find all the information. http://www.gils.net/.

GMES. 'Global Monitoring for Environment and Security' (GMES). *http://www.gmes.info/*

INSPIRE. Infrastructure for Spatial Information in the European Community. http://inspire.jrc.it/

INSPIRE. (2008). Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32008R1205:EN:NOT

ISO 19115. Geographic information – Metadata. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=26020

LinkingOpenData. W3C SWEO Community Projekt. http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData

OAEI. (2007). Ontology Alignment Evaluation Initiative. http://www.few.vu.nl/~wrvhage/oaei2007/environment.html

OGC CAT. OpenGIS Catalogue Service Implementation Specification. http://www.opengeospatial.org/standards/cat

Rüther, Maria and Thomas Bandholtz. Semantic Network Service (SNS) Update and Future Vision. EnviroInfo, September 10-12, 2008, Leuphana University of Lüneburg, Germany. http://bandholtz.eu/publications/2008-SNS-EnviroInfo08.pdf

SEMIC.EU The Semantic Interoperability Centre Europe. http://www.semic.eu/

UBA. (2008), Scientific Library for Environmental Protection of the Federal Environment Agency. http://www.umweltbundesamt.de/uba-info-e/e-fabib.htm

UMTHES. (2008). Umwelt-Thesaurus - Terminologie zu Umweltnutzung / Umweltbelastung / Umweltschutz. http://www.umweltbundesamt.de/uba-info/dokufabib/thes.htm

URI. (1998/2005). Uniform Resource Identifiers (URI): Generic Syntax. IETF RFC 2396. August 1998. Current Version: RFC 3986. 2005. http://www.rfc-editor.org/rfc/rfc3986.txt

W3C. (2008). Cool URIs for the Semantic Web. W3C Working Draft 21 March 2008. *W3C Interest Group Note 31 March 2008*. http://www.w3.org/TR/cooluris/

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Discovery and Analysis of Environmental Information based on Formalised Terminology

U. Bügel[a], **T. Usländer**[a]

[a] *Fraunhofer IITB, Fraunhoferstr. 1, 76131 Karlsruhe*
(ulrich.buegel|thomas.uslaender@iitb.fraunhofer.de)

Abstract: Although the use of formalised environmental terminology nowadays gets popular in practice, the full benefit that can be drawn from it them is by far not yet reached. In terms of expressivity, the most advanced form of environmental terminologies are ontologies which are specified formally by means of formal logics. The availability of formal ontologies promotes automation of content provision and integrated access to structured and unstructured information resources via customised user interfaces for domain experts and public users. The paper introduces mechanisms for semantic annotation and automatic classification as essential base technologies that open the door to better information discovery, analysis and integration. We report experiences of their application in selected projects, namely the development of semantic extensions for geospatial catalogues in the context of the ORCHESTRA project and ontology-based information integration of environmental information systems in Baden-Württemberg/Germany. Thereby, special attention is paid to the fact that semantic extensions exert a minimum of impact to the available, grown infrastructure for conventional information retrieval.

Keywords: Environmental Terminology; Semantic Annotation; Ontology Based Ranking; Automatic Classification; Semantic Catalogue; ORCHESTRA.

1. INTRODUCTION

Nowadays, the value of environmental terminologies as a basic asset for improvement of information retrieval in Environmental Information Systems (EIS) is unquestionable. However, transfer into practice has not reached a satisfying stage. Not all benefits are yet fully exploited. Although environmental portals such as the Semantic Network Services (Bandholtz [2003]) in the German PortalU (http://www.portalu.de/) allow integrated access to data and services through a common terminology, considerable challenges in the fields of automated content provision, integration of structured and unstructured information, exploitation of advanced descriptive techniques and customisable user interfaces are still evident.

EIS are typically either directed towards thematic experts providing highly detailed information and sophisticated user interfaces, or, mostly in a more restricted and aggregated form, to the public user, e.g. the citizen who aims at getting environmental information in his surroundings. However, the need to fulfil environmental directives on regional, national and international level increasingly leads to mix up these user communities. The major technological challenge is to provide proper interfaces for search, navigation and analysis to so-called "inexperienced" users.

The development of terminologies is influenced by technological progress and observed trends. The most advanced specification techniques are developed in the context of the Semantic Web initiative (Berners Lee [2001]). The World Wide Web Consortium (W3C) has approved standards by which ontologies can be specified in terms of formal logics with high expressivity. Our approach starts from the assumption that a common environmental terminology is available and formally specified in terms of an ontology. We then describe how technologies such as ontology based semantic annotation and automatic classification can leverage the explication of formal knowledge that is implicitly present in content, and how they can foster solutions with a high degree of automation. In the sequel, approaches to intelligent discovery, analysis and integration of information are outlined.

An observed problem with commissioning of Semantic Web technology is that it often tends to disregard the presence of an available solid and grown infrastructure for information retrieval. In order to achieve acceptance by organisations that operate an EIS, we assume that components of their deployed IT infrastructure (e.g. catalogues, search engines) should not be replaced, but remain operational. We elaborate this approach by illustrating some ideas and scenarios found in various projects.

2. TECHNOLOGICAL IMPACTS

2.1 Ontologies

An ontology provides system independent description of knowledge. It represents a shared and agreed view of domain experts on a certain part of the real world in a standardised language dedicated to machine processing. According to Smith & Welty [2001] ontological descriptions can be classified according to their expressiveness. The range starts from simple structures as found in glossaries up to more complex structures such as common hierarchies of terms. Additional relations between terms such as synonyms can be found in standardised Thesauri. In the field of the Semantic Web ontologies are conceived as taxonomies described by a set of classes ("concepts") which are specified as objects with properties (e.g. "name", "description"). Properties of more general classes are inherited by the more specific ones. Moreover, arbitrary relationships between classes can be defined (e.g. "person is_employed_at company"). The highest expressivity can be achieved by describing the classes, properties and instances by a set of logical constraints, axioms and integrity conditions. The most advanced standard for ontologies used in the Semantic Web is OWL (Smith M.K [2004]).

2.2 Semantic Annotation

In the Semantic Web a basic idea is to provide *semantic annotation* of content (and also services, databases etc.) with formal meta-information which can be embedded into a commonly understood context – given by means of ontologies – and further processed by computer applications. While today most tools for semantic annotation are highly interactive, the European Integrated Project ORCHESTRA (see Usländer et al. [2007]) has developed an annotation service which can *automatically* generate meta-information from various sources and relate it to elements of an ontology (e.g. concepts properties, instances) (see Kopp [2007]). Annotation of unstructured and semi-structured sources (e.g. texts in natural language, documents or web pages) is based on a tool for automatic information extraction Cunningham et al. [2007]). This tool is able to identify and normalise named entities (e.g. rivers, cities) that occur in the text by utilizing methods of Natural Language Processing (NLP). Annotated text passages can be highlighted and their semantic bounding box can be explored such that these passages can be interpreted by means of the environmental

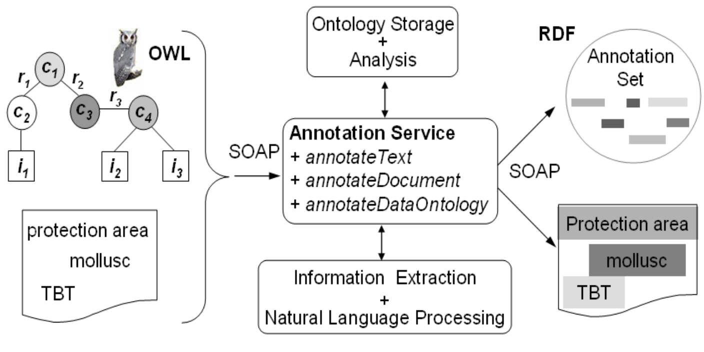


Figure 1: The ORCHESTRA Annotation Service

terminology. Entity identification in texts is based on pre-populated knowledge stored in a knowledge base, e.g. simple facts such as “the Rhine is a river”. Currently, the service is further developed in order to cover detection of instances of object properties defined in the ontology. For example, a property “chemicals leach_into water” could have an instance “TBT leaches_into North_Sea” which can be detected in texts such as “… however, TBT does not remain at the ships surface, it leaches further into the North Sea …”.

2.3 Automatic Classification

While semantic annotation is the basis for the generation of formalised knowledge stored in knowledge bases and retrieved by means of semantic search user interfaces, *automatic classification* is the means for assigning content to a navigational structure. It disburdens the user from the task of manual classification of content to be inserted.

Automatic classification means assignment of objects to categories according to their properties. Classification of texts or documents may be realised through various approaches. If the categories are given through a formalised terminology, e.g. an OWL ontology, each document can be assigned to the best fitting ontological concept. The classification can be calculated through semantic annotation of the document followed by application of weighting algorithms on the annotations. Often the categories are defined much simpler, e.g. by a notation of the hierarchical term structure or by means of a given hierarchical document store with named folders and some documents already inserted manually. In such a scenario, the application of methods for supervised learning - a Machine Learning approach - is a proper choice for the accomplishment of the classification task. Here, the documents to be classified are compared with available documents which have previously been classified manually by domain experts. These documents and the pre-classifications build a *training corpus* from which a classifier can be generated that is able to classify further documents automatically. Figure 2 outlines the steps to be performed in a supervised learning process.

After the training corpus has been composed, all corpus documents are to be transformed into a representation form for machine processing. From each document, a term list is extracted (so-called “features”), which contains all terms that can be used for a classification of the document. This is achieved by means of a text processing software which executes a linguistic and a statistic analysis of the text. Thereafter, each training document – and also each document to be classified automatically – is transferred into a feature representation form. It consists of a vector with weighted features occurring in the document. For the weighting, a spectrum of methods is available. For instance, the method *Term-Frequency * Inverted Document Frequency (TFIDF)* weights the frequency of each document term and decreases the weighting for each occurrence of the term in other documents. Two feature vectors can now be compared by means of a similarity function, e.g. the cosine of the angle between the vectors. This enables the definition of an automatic classifier. For large amounts of texts the utilization of support vector machines as classifiers has shown the best results (Joachims [2002]).

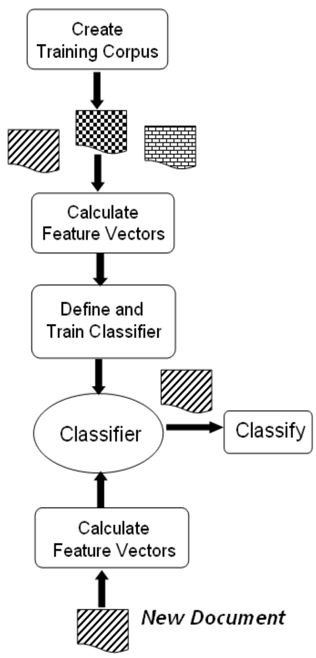


Figure 2: Automatic Classification

The architecture shown in figure 2 has been implemented as a Web service (Frank, [2008]). For the composition of the corpora, the linguistic analysis, the feature calculation and the classification task the data mining workbench “RapidMiner” (Mierswa [2006]) has been used, which provides also for selection of various classifiers such as Bayes classifier, decision trees or support vector machines.

3. ONTOLOGY BASED DISCOVERY AND NAVIGATION

In this section, we illustrate how formalised terminology is used advantageously in projects for the design of search and navigation interfaces, especially for inexperienced users.

The geospatial community has developed standard services called "(geospatial) catalogues" by which information about resources, being information or services, can be found. Geospatial catalogue services provide access to (meta-)information about available geo-spatial resources, e.g. topographical data or geo-statistical processing capabilities. Widely used for this type of application are catalogues which are compliant to specifications of the Open Geospatial Consortium (OGC).

Inexperienced users of such catalogues do often not know the correct search keywords that exactly match to stored meta-information which may result in a poor recall. In ORCHESTRA a semantic catalogue has been defined and implemented which offers semantic extensions while still providing a conventional catalogue search interface (Hilbring and Usländer [2008]).

Keywords contained in queries are used to identify corresponding concepts defined in an uploaded ontology. Starting from a concept, other concepts that reside in a kind of "semantic bounding box" determined by ontological relationships (parents, children and property relationships) are additionally looked up in the ontology. The catalogue expands the terms of the query according to the depth of the semantic bounding box and sends individual queries for each of the concepts found. The user obtains an easy-to-use search interface: by extending or narrowing the bounding box he is able to navigate through the existing environmental terminology before he exactly formulates his search query.

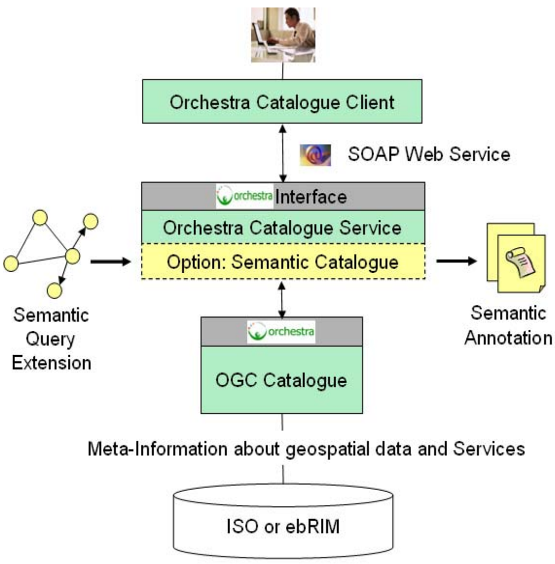


Figure 3: Semantic Extensions in ORCHESTRA Catalogues

A further issue to be considered is the quality of the results of search queries. With current search technology, users with a concrete objective in mind often need to submit series of requests to (one or several) EIS and collect all retrieved information which is relevant to take a decision. Consider a citizen who wants to get an overview about environmental conditions (e.g. transport connections, noise disturbance, air quality, land contamination) in a given geographical area. By using available portals for environmental information retrieval the user faces several problems here. Due to a lack of integration of thematically different systems he gets unsatisfying results because combination of thematically different search items often does not return the desired information. Moreover, integration of geospatial information to the thematic search items is not given. As a consequence of this, the user has to retrieve lots of documents with different search terms, compose the results manually together and somehow verify that he has got the complete information together.

An ontology specifying all important relationships of a concept "environmental conditions" can be the basis for automatic generation of the complete context when a single query for "environmental conditions" is issued.

In a recently launched project of the German environmental cooperation network KEWA [Mayer-Föll et al, 2009] such an ontology is being developed and applied for an integrated search over environmental systems in Baden-Württemberg, the content of which is semantically annotated with the entire ontology. An integrated search is investigated by several approaches. For instance, a result entry of a search request "environmental conditions at …" displays the geographical area of the location and is automatically accompanied with generated links to all contents of interest (noise disturbance, air quality etc.), derived from the entire ontology. A further possibility is to exert influence on the ranking of search results by means of the ontology, i.e. the search results are ordered according to their semantic closeness to the search item. For instance, entries about "air quality" are ranked higher despite to the fact that the terms "air" and "quality" have not been search terms. The implementation of the ranking issue is more elaborated in the next section.

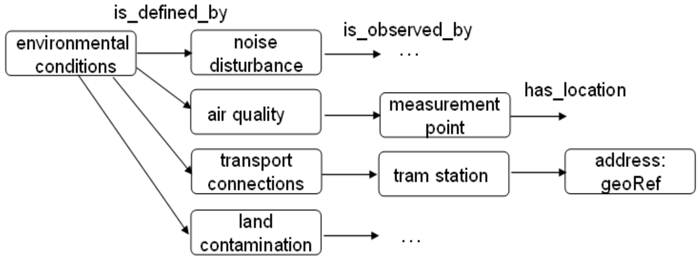


Figure 4: Excerpt of an Ontology „Environmental Conditions"

3. ANALYSIS AND RANKING OF SEARCH RESULTS

3.1 Ontology Based Ranking

The ORCHESTRA Semantic Catalogue uses the ontology also for an optional search extension – the ranking of search results. Currently the following simple ranking algorithm is implemented. Firstly, the ranking extension identifies semantically related keywords to literals contained in the original catalogue query. Secondly, it searches for newly identified keywords in the conventional catalogue search result and calculates weightings for hits. Thirdly, the search result is sorted according to the assigned weightings. Finally the ontology based ranked result is given back by the search operation.

The semantic extensions for OGC catalogues can be configured as an option that provides additional functionality. They have been realised on top of existing technology and cause only little impact on the entire operating infrastructure. The implemented ranking algorithm directly operates on the results received by queries to the underlying native OGC Catalogue(s). In the ontology development project for environmental systems in Baden-Württemberg - outlined in section 2 – a similar approach with smooth integration into the grown infrastructure is pursued (Ebel et al. [2008]). These systems utilize the Google™ Search Appliance (GSA), a search engine with very good performance. Any approach for semantic extensions starting from scratch and introducing new technology replacing the existing one would be unacceptable. However, considering the ontology based ranking extensions, a different situation compared to the Semantic Catalogue is given: Search results delivered by a Google engine may not be reordered due to company policy. Therefore, another implementation approach is taken.

Figure 5 outlines the architecture. The GSA crawls for documents located at Web pages of the environmental information systems in the domain and builds a search index that provides a high-performance response to user queries. The idea to take influence on the ranking of the search results by means of the ontology is to semantically annotate the documents and store the annotations in the meta tags of the Web page on which they reside. For instance, pages (or documents on pages) containing information about air quality, traffic connections, contamination etc. are annotated with "environmental conditions" as specified in the ontology by a respective relation. The information stored in the meta tags thus will be included in the search index of the GSA. A configuration parameter of GSA "one-boxes" enforces to weight information contained in meta tags higher in the search index. This mechanism is applied in order to enforce that these pages appear as top hits in the results of a search for "environmental conditions", despite to the fact that the search terms eventually are not contained in the page. From a users point of view, the described method is a "ranking by annotation" similar as in the catalogue example. However, the annotations here are not executed on search results as with the catalogues, but on entries which are characterised with concepts and relations defined in the ontology. This is done in a way such that the underlying search engine takes these annotations into account when building the search index and automatically ranks higher those entries with semantic closeness to the search terms. Note that the mechanism does not imply any re-ranking of the results of an underlying conventional search engine.

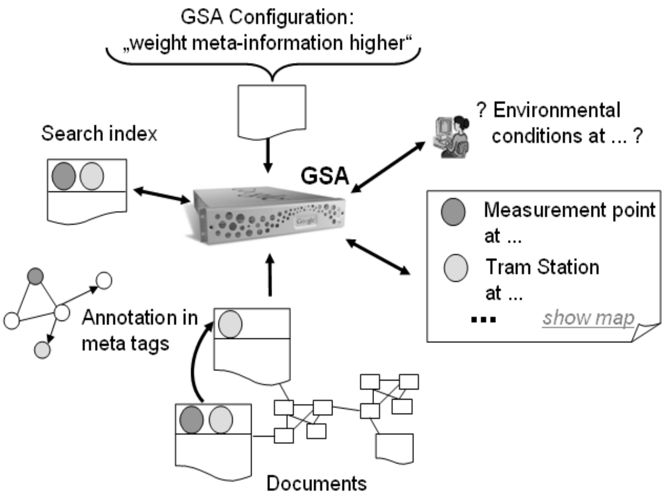


Figure 5: Ontology Based Ranking in a Google™ Environment

3.2 Interpretation of Discovered Content

When looking at the results of a search query, inexperienced users may face another problem. Texts and data obtained through queries are often not self-explanatory, and users need assistance in order to understand them. A given environmental terminology can lay the foundation for a "reading assistance" based on semantic annotation. A client of the ORCHESTRA Annotation Service is able to highlight text passages and explore their semantic bounding box such that these passages can be interpreted by means of the environmental terminology (figure 6).

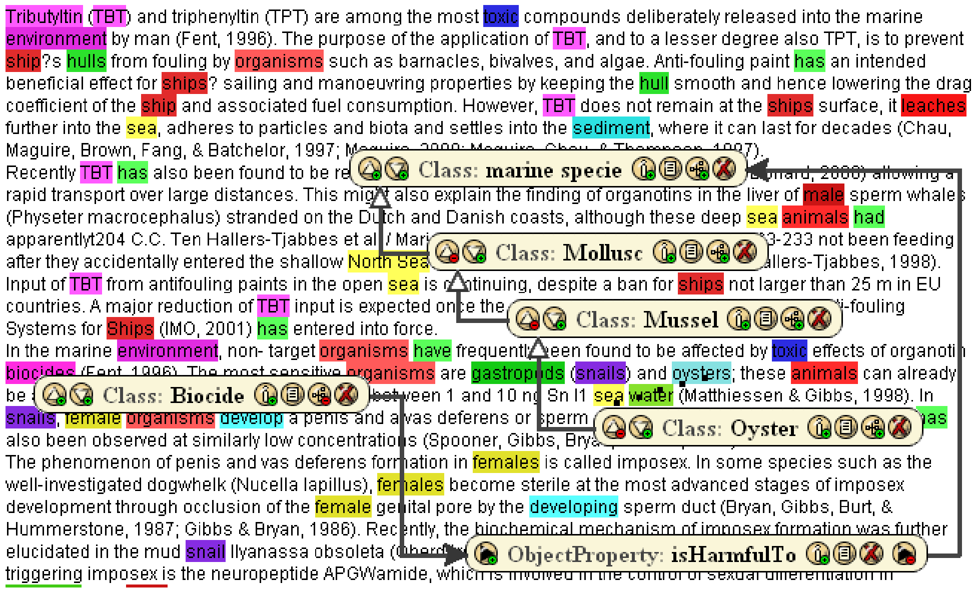


Figure 6: Semantic Annotation of Text Based Search Results

The user obtains a comprehensive assistance for interpretation of the text. Moreover, he has the possibility to redirect identified concepts shown in the spawned ontology browser to the interface of the catalogue client and start a new search for information.

4. CONNECTING STRUCTURED AND UNSTRUCTURED INFORMATION

In order to perform tasks in the area of environmental protection access to (geospatial) information held at other administrations such as geology, forestry, agriculture, urban and land use planning, statistics, traffic etc. is needed. This information often is available as structured data in environmental databases. For geodata networks such as WIBAS (Baden-Württemberg) and integrated information systems such as ALKIS/ATKIS, so-called feature type catalogues are being defined for all domain areas which represent those features that describe real word phenomena. This facilitates the data exchange.

In addition to the structured information a considerable amount of unstructured information in form of texts, documents etc. is held in domain information systems, e.g. leaflets, technical bulletins, executive regulations or background information related with the data. In many cases, these documents have an implicit semantic relationship to concrete objects in the feature type catalogues. When retrieving environmental information it is helpful to link structured information in databases with unstructured information contained in documents. Therefore, these relationships have to be explicated. Here, semantic annotation and automatic classification can significantly help. This is illustrated in two examples.

Consider a user is reading a document containing a text passage about a certain measurement point (figure 7). When reading the text, it may be of value to the user if observations of this measurement point could be directly requested and displayed as additional information. This can be achieved by annotating the text and use the annotations for automatic formulation of a proper database query. In order to avoid a too detailed presentation to the user, the instance description of the measurement point in the ontology already contains selectors for attributes to be displayed. These attributes are inserted into the query. Thus, the reporting system of the database may exactly generate the desired view on the data.

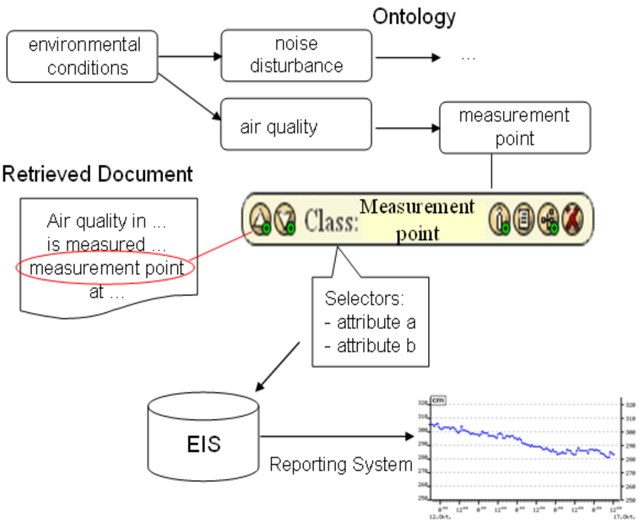


Figure 7: EIS Access from Annotated Documents

Such a linkage can also be useful in the opposite direction. If the user has retrieved information from a database about a measurement point, he may be interested to look at accompanying information of interest contained in documents. A possible solution to that is to classify documents according to the application schema of the feature catalogues (figure 8). The feature catalogue is given here as an ontology. The documents are stored in an “intelligent store” where each document has been assigned to the ontological concept representing its best possible

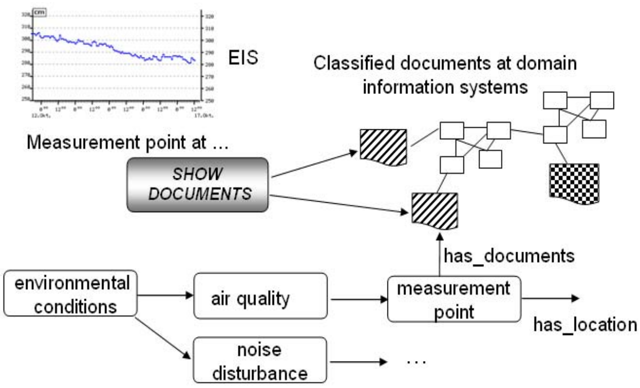


Figure 8: Retrieval of Accompanying Documents

classification. The classifications are accomplished by a semi-automatic process, where a trained classifier makes classification suggestions and a human expert confirms or rejects them. When information about a measurement point retrieved from a database is displayed, a query to the ontology browser can be constructed who generates and displays a list of references to all documents with the proper classification.

6. CONCLUSIONS

Although considerable research work is being performed in the domain of geospatial semantics, the main driving force for the integration of semantic technologies into service-oriented architectures are the Semantic Web initiative and the related W3C standardisation efforts. However, technologies emerging in the Semantic Web area still have some academic flavour and their integration into grown infrastructures for conventional information retrieval is a challenge. The projects described in this paper contribute to the transfer of semantic technologies into practice according to user requirements and to foster best exploitation of environmental information technologies. Finally, it can be stated that semantic terminologies can help to open the door for inexperienced users (e.g. citizens) to information contained in EIS originally designed to be used by domain experts. Reuse of an existing EIS infrastructure requires careful selection, adaptation and operation of available semantic tools in order to be accepted by the EIS community.

ACKNOWLEDGEMENTS

This work has been supported by the European Projects ORCHESTRA co-financed by the the European Commission in its FP6 IST programme, and the environmental research network KEWA led by the Ministry of the Environment Baden-Württemberg, Germany.

REFERENCES

Bandholtz, T. Implementation of a Semantic Network Service (SNS) in the context of the German Environmental Information Network (gein®), http://www.semantic-network.de/sns-summary-2003-06-06.pdf, 2003.

Berners Lee, T., Hendler, J., Lassila, O., The Semantic Web, *Scientific American*, May 2001

Cunningham, H., Maynard, D., Bontcheva, K., Tablan, V., Developing Language Processing Components with GATE Version 4 (a User Guide), 2007.

Ebel, R. et al. Die Google Search Appliance und das Umweltinformationssystem Baden-Württemberg, http://www.google.de/intl/de/enterprise/case_studies/lubw.pdf, 2008.

Frank, T., Entwurf und Implementierung eines Dienstes zur automatischen Klassifikation von Texten. Diploma Thesis, TU of Berlin / Fraunhofer IITB, Karlsruhe, 2008.

Hilbring, D., Usländer, T., Ontology-Based Discovery of Geoscientific Information and Services. *Session on Semantic Interoperability, Knowledge and Ontologies in the European Gesciences Union General Assembly in Vienna*, 13-18 April 2008

Joachims, T., Learning to classify text using support verctor machines: methods, theory, and algorithms. Kluwer Academic Publishers, Boston/Dordrecht/London, 2002.

Kopp, A., Design and Implementation of an Automatic Semantic Annotation Service, Diploma Thesis, Universität des Saarlandes / Fraunhofer IITB, Karlsruhe, 2007

Mayer-Föll, R., Keitel, A., Geiger, W. (Eds.). F+E Vorhaben KEWA Phase IV 2008/09, http://www.lubw.baden-wuerttemberg.de/servlet/is/27023/, 2009.

Mierswa, I., Wurst, M., Klinkenberg, R., Scholz, M., Euler, T., YALE: Rapid Prototyping for Complex Data Mining Tasks. *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-06*), 2006.

Smith, M. K., Welty, C., McGuinness, D. L., OWL Web Ontology Language Guide, http://www.w3.org/TR/owl-guide, 2004.

Smith, B., Welty, C., Ontology: Towards a New Synthesis. *Proc. of the International Conference on Formal Ontology in Information Systems (FOIS2001)*. ACM Press, 2001

Usländer, T. (Ed.), RM-OA - Reference Model for the ORCHESTRA Architecture V2. *OGC Best Practices Paper 07-097*, http://www.opengeospatial.org/standards/bp, 2007.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Environmental Thesaurus UMTHES – environment terminology evolvement for a broad thematic matter and for different use conditions – practicably handling of structural challenges

J. Fock

Environment Protection Agency, Dessau, Germany (joachim.fock@uba.de)

Abstract: The development of the environment terminology within "Environmental Thesaurus" (Umweltthesaurus UMTHES, Federal Environment Agency Germany = Umweltbundesamt Deutschland) and adaptation to changing European information demands (INSPIRE). Generating INSPIRE-keywords (INSPIRE-themes) is described.

Keywords: *Environmental terminology; semantics; Environmental Thesaurus; Semantic Network Service; SNS; INSPIRE*

1. INTRODUCTION

European environmental information shall become accessible all over the European Community and between the member countries. Not at last European institutions want to get direct access to environmental information in member states. For this, projects like INSPIRE (Infrastructure for Spatial Information in Europe: "aiming to assist policy-making in relation to policies and activities that may have a direct or indirect impact on the environment") and SEIS (Shared Environmental Information System: "… Boosting intelligent eGovernment services for the Environment") are developed. Naturally the topic ENVIRONMENT is strongly interdisciplinary.

This paper describes the approach to serving demands from this development by means of Environmental Thesaurus UMTHES. UMTHES is the semantic basis for the Semantic Network Service SNS, which is the German instrument for keyword metadata generation of the German national environment information portal (PortalU) and different other environmental information sources e.g. the Environmental Research Database (Umweltforschungsdatenbank UFORDAT). UMTHES includes nearly all of the GEMET terminology in English and German language (General Environmental Multilingual European Thesaurus) and represents the semantic reservoir to serve the demands of INSPIRE and SEIS for the keyword metadata.

2. UMTHES

2.1 History and development

The "Umwelt-Thesaurus" (UMTHES® =environmental thesaurus) had been founded in about 1976 as a systematic creation.

When starting real use of UMTHES® (UMweltTHESaurus = environmental thesaurus) in the 1980's the application fields was straitened to two databases only: the environmental literature data base ULIDAT (=Environmental-Literature Database) and the environmental research data base UFORDAT (=Environmental-Research Database). The subject of the databases was defined by the so called "Umweltklassifikation" (Environmental Classification Scheme) (Table 1).

Table 1 Environmental classification scheme

AB	Abfall	**Waste**
BO	Boden	**Soils**
CH	Chemikalien/Schadstoffe	**Environmental chemicals / pollutants**
EN	Energie- und Rohstoffressourcen - Nutzung und Erhaltung	**Environmental aspects of energy and raw materials**
GT	Umweltaspekte gentechnisch veränderter Organismen und Viren	**Environmental aspects of organisms and viruses, modified by genetic engineering**
LE	Lärm / Erschütterungen	**Noise / Vibrations**
LF	Umweltaspekte in Land- und Forstwirtschaft, Fischerei, Ernährung	**Environmental aspects of agriculture, forest and fishing, food**
LU	Luft	**Air**
NL	Natur und Landschaft/Räumliche Aspekte von Landschaftsnutzung, Siedlungs- und Verkehrswesen, Tourismus und urbaner Umwelt	**Nature and landscape / Regional Development**
SR	Strahlung	**Radiation**
UA	Allgemeine und übergreifende Umweltfragen	**General and interrelated environmental topi**
UR	Umweltrecht	**Environmental law**
UW	Umweltökonomie	**Environmental economics**
WA	Wasser und Gewässer	**Water and Waters**

This Environmental Classification Scheme already covers all aspects of "environmental information" as later defined by DIRECTIVE 2003/4/EC (… public access to environmental information …).
Geographically the data bases focussed to publications and R&D-projects from German-speaking Europe. Literature from mainly US and UK was not includes because of existing American databases ENVIROLINE and POLLUTION. Hence the development started monolingual only in German.
When it had become in use, the thesaurus was extended with keyword concepts from documented literature and research descriptions. This is still one way of continuing thesaurus development further on.
English language also in literature on environmental topics became more important. So we enriched the thesaurus with English equivalents. At least every preferred term should be translated. Today UMTHES has about 30,000 English entries.

2.2 UMTHES today

UMTHES thesaurus today offers

- Term name
- Term stem
- Different spellings
- The equivalence relationship
- The hierarchical relationship
- The associative relationship
- English equivalence as individual, adequate thesaurus entries
- Morphological information
- Definitions, usage notes, source reference, GEMET-reference
- Assignment of concepts to environmental classification, INSPIRE themes,
- Possible paraphrases for text analysis purposes

The sources for terminology development:

- Literature
- Log files
- Other Thesauri
- Glossaries…
- WEB-Retrieval
- Regulations, first of all EU-regulations

Indexing from thesaurus is combined with use of the above mentioned environmental classification scheme where possible. Documents content is indexed from the thesaurus and thematically classified with the classification. The environmental classes do relate the subscripted documents to certain environmental topic. This has big importance for the quality of retrieval. A term, e.g. "emission", could get another, often more special meaning when the document as a hole is classified to different classes, e.g. to the class "Noise" or "Air"(noise emission vs. air pollution).
For semi-automatic suggestion of document classes, the thesaurus terms (preferred terms) have been assigned to environmental classes of the classification scheme. Via occurrence and statistical analysis proposals for document classification assignments could be generated. This experience led to the recent generation of INSPIRE-Themes proposals from and for metainformation of INSPIRE-object in German catalogues of data sources (UDK=Umweltdatenkataloge) (see below).
Today, in principle this practice for disambiguation has some similarity to the use of INSPIRE themes or the topic categories: "The topic category is a high-level classification scheme to assist in the grouping and topic-based search of available spatial data resources … While the topic category is too coarse for detailed queries, keywords help narrowing a full text search and they allow for structured keyword search." (COMMISSION REGULATION (EC) No 1205/2008, Part B Metadata elements 3.1)

But the INSPIRE-themes and "topic categories" are more focused on spatial information aspects in contrast to global view of environmental classification scheme for "environment". May be, the broader view will become interesting again when establishing SEIS because it will not be restricted on spatial information. By its definition of environmental information DIRECTIVE 2003/4/EC (… public access to environmental information …) offers a broad themes-catalogue. It already has been implemented in German environmental internet portal "PortalU" for thematic structuring and information retrieval.

In recent time the retrieval of information is done by information end users. They search in the data base or internet search engines by using screen query forms and not longer by structured query languages. The users ordinarily must not know anything about the structures of the searched data bases. Searching of preferred term keywords from lists is not reasonable today, so hints to necessary disambiguation by parenthetical qualifiers e.g. are not attended. Familiarity with the problems of homonyms can not be assumed. As a solution automatic text analysis is used for the analysis of entries of query search forms. Natural language requests are tried to be effectuated by enrichment of thesaurus with naturally spoken phrases. Complex terms or compound terms from the thesaurus can be determined by featured software. From typed in searching words and paraphrases on basis of complex thesaurus terms in many cases disambiguation may be obtained. Such entries have a part in doubling non-preferred terms in UMTHES thesaurus to 58,000 today.

2.3 UMTHES in SNS

Today, a main user of SNS, the German Environmental Information Portal (=Umweltportal Deutschland http://www.portalu.de) offers central access to over 1,000,000 web pages and about 500,000 database entries from public organisations in Germany.
These pages from the Internet are very different or poor structured. No reliable document area can be shown as a rule for successful text analysis. Common words from anywhere in such pages may have very different meaning. Considering such common words only in cases where potent qualifying words are adjacent, may be a method for disambiguation. For that reason useful term combinations are introduced into the UMTHES thesaurus. For the

retrieval software such possibilities of disambiguation on base of the thesaurus feature are eligible for future.

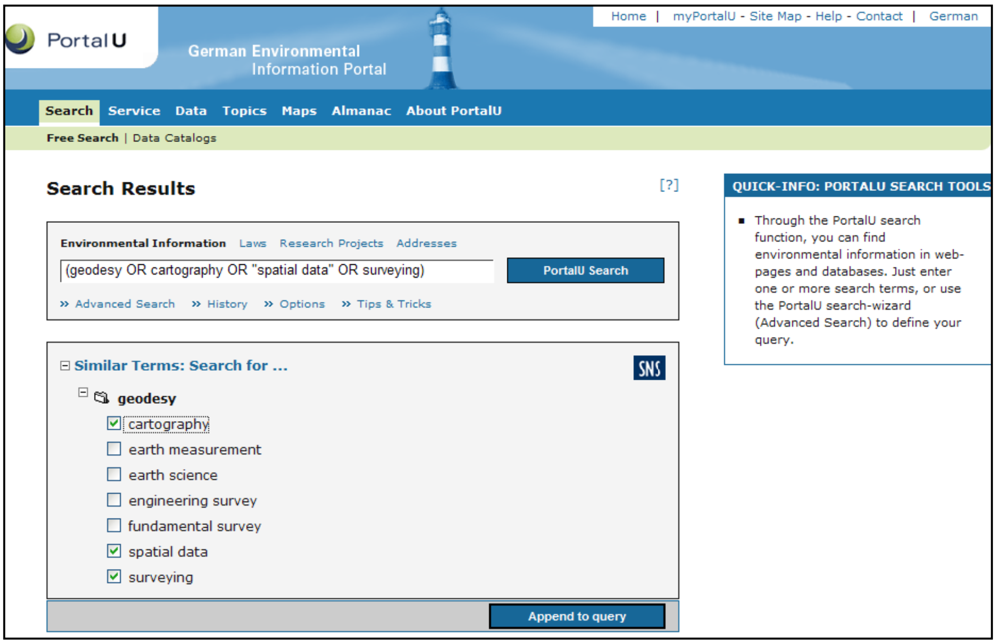


Figure 1: supplementary search terms, "similar terms", from SNS

Public use of the environmental terminology UMTHES is offered by Semantic Network Service (SNS) (Semantic Net Service=Semantischer Netz Service).

SNS is a bilingual (German/English) semantic network which consists of three components:

- The Environmental Thesaurus UMTHES® with more than 50,000 inter-networked terms. (Descriptors (preferred terms) and Non-Descriptors).
- The Geo-Thesaurus-Environment (GTU) with more than 25,000 geographic names and the spatial intersections of all these places.
- an Environmental Chronology containing more than 600 contemporary and historical events that affected the environment

The Federal Environment Agency provides this service to encourage the practice of a common domain language in the German Environmental Informatics, and to make this terminology accessible for the public.

Services within SNS include

- All about a Topic. **getPSI**
- Search for Topics. **findTopics**
- Search the Chronology. **findEvents**
- X Years ago .... **anniversary**
- Word Choice Support. **getSimilarTerms**
- Sub-Trees in a Snap. **getHierarchy**
- Indexing of URL . **autoClassify**
- Indexing of Text . **autoClassify**

2.4 UMTHES and INSPIRE

Demands of the INSPIRE-Directive ("Infrastructure for Spatial Information in EUROPE")

Background: "Whereas: … (4) The value domain of each metadata element (in INSPIRE) is necessary to ensure interoperability of metadata in a multilingual context and that value domain should be able to take the form of free text, dates, codes derived from international standards, such as language codes, keywords derived from controlled lists or thesauri, or character strings (COMMISSION REGULATION (EC) No 1205/2008, Preface)

DIRECTIVE 2007/2/EC - CHAPTER IV - NETWORK SERVICES Article 11:

1. Member States shall establish and operate a network of the following services for the spatial data sets and services for which metadata have been created in accordance with this Directive:

(a) discovery services making it possible to search for spatial data sets and services on the basis of the content of the corresponding metadata and to display the content of the metadata;

(b) ...

2. for the purposes of the services referred to in point (a) of paragraph 1, as a minimum the following combination of search criteria shall be implemented:

(a) keywords

(b) ...

Of the requirements of INSPIRE concerning metadata, above all, the following are of concern for UMTHES:

- Keywords derived from controlled lists or thesauri – "3. KEYWORD : ... If a resource is a spatial data set or spatial data set series, at least one keyword shall be provided from the general environmental multilingual thesaurus (GEMET) describing the relevant spatial data theme as defined in Annex I, II or III to Directive 2007/2/EC. For each keyword, the following metadata elements shall be provided: 3.1. Keyword value: The keyword value is a commonly used word, formalised word or phrase used to describe the subject. While the topic category is too coarse for detailed queries, keywords help narrowing a full text search and they allow for structured keyword search.(underlined by author)
- Topic category - The topic category is a high-level classification scheme to assist in the grouping and topic-based search of available spatial data resources" (TOPIC CATEGORIES IN ACCORDANCE WITH EN ISO 19115 - COMMISSION REGULATION (EC) No 1205/2008)

UMTHES features assignments between "Spatial Data Themes" (INSPIRE-Directive, Annex I, II, III), and German equivalents "Geodatenthemen" of German "GeodatenZugangsGesetz - GeoZG" and the assignment between "Topic Categories" and "Spatial Data Themes" according to COMMISSION. REGULATION (EC) No 1205/2008 (Part D, 2.).

GEMET-terms are part of the environmental thesaurus UMTHES (both in English and German language, all about 10,000). Terms with special closeness to certain Spatial Data Themes are assigned in UMTHES to these themes.

With help of this allocation we now want to support the metadata processing of the Environmental Data Catalogue (Umweltdatenkatalog UDK) datasets for the provision to INSPIRE.

In the German Environmental Data Catalogues within the German Environmental Information Portal "PortalU®" a tagging with UMTHES thesaurus keywords is carried out. Spatial Data Themes could be assigned to many of these UMTHES keywords. From this we want to generate suggestions for marking datasets with Spatial Data Themes. If keywords point to certain Spatial Data Themes, these Spatial Data Themes are shown for selection in a choice list.

Vice versa a list of search terms can be generated for discovering information about an interesting Spatial Data Theme. Because additional Keywords should be attached, This would be a useful support for metadata generating for INSPIRE:

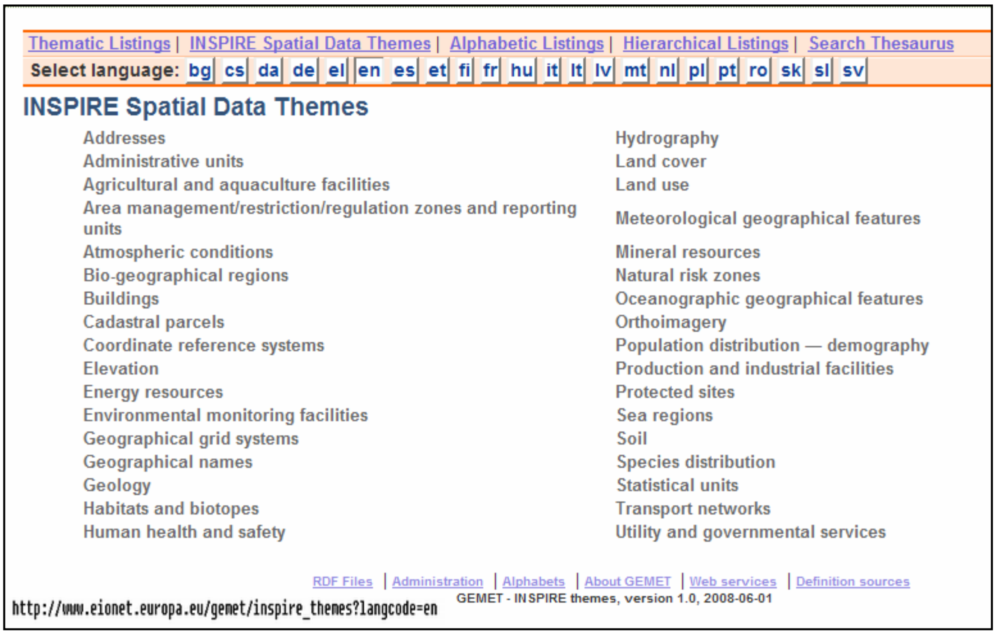


Figure 2: GEMET - integrated INSPIRE spatial data themes

"According to the INSPIRE Implementing Rule for Metadata, if a resource is a spatial data set or spatial data set series, at least one keyword shall be provided from the General Environmental Multi-lingual Thesaurus (GEMET) describing the relevant spatial data theme as defined in Annex I, II and III to Directive 2007/2/EC. The titles and definitions of all 34 INSPIRE Spatial Data Themes have now been integrated into GEMET" For the improvement of the information discovery additional Keywords should be attached." (INSPIRE Metadata Editor User Guide 2008-05-15)

The 34 Spatial data themes of Annex I, II and III of INSPIRE Directive are an essential selection criteria for spatial data sets becoming subject of the INSPIRE Directive.

- Article 4 1: This Directive shall cover spatial data sets which fulfil the following conditions: ... (d) they relate to one or more of the themes listed in Annex I, II or III (INSPIRE Directive 2007)

Annex I contains 9 themes e.g. "Coordinate reference systems" and "Protected sites"

Annex II contains 4 themes e.g. "Elevation" and "Geology"

Annex III contains 21 Themes e.g. "Statistical units" and "Mineral resources"

For every data set it has to be checked, whether it covers one of the themes and therefore is covered by INSPIRE-regulations. Probably the persons in charge may decide this by feeling.

Moreover – for spatial data sets and for spatial data sets series the INSPIRE Implementing rules for metadata (COMMISSION REGULATION (EC) No 1205/2008) mandate for metadata of spatial data sets … the presence of at least one keyword, describing the relevant INSPIRE spatial data theme (as defined in Annex I, II and III of the INSPIRE Directive) originating from the general environmental thesaurus (GEMET).

Obviously meant is one or more of the themes names, as they became integrated into GEMET in 2008. Figure 2 gives an overview on the INSPIRE Spatial Data Themes.

From experience it seems to be more difficult to decide which of the 34 themes is or are exact convenient to the thematic coverage of a single data set or …series content.

We try to facilitate this by assigning the INSPIRE-themes to selected concept in UMTHES. Using the Semantic network Service (SNS)

In a first step we assigned just 85 broader-term concepts of UMTHES to INSPIRE-themes. Integration of narrower terms makes this tool much more powerful, as Figure 3 shows.

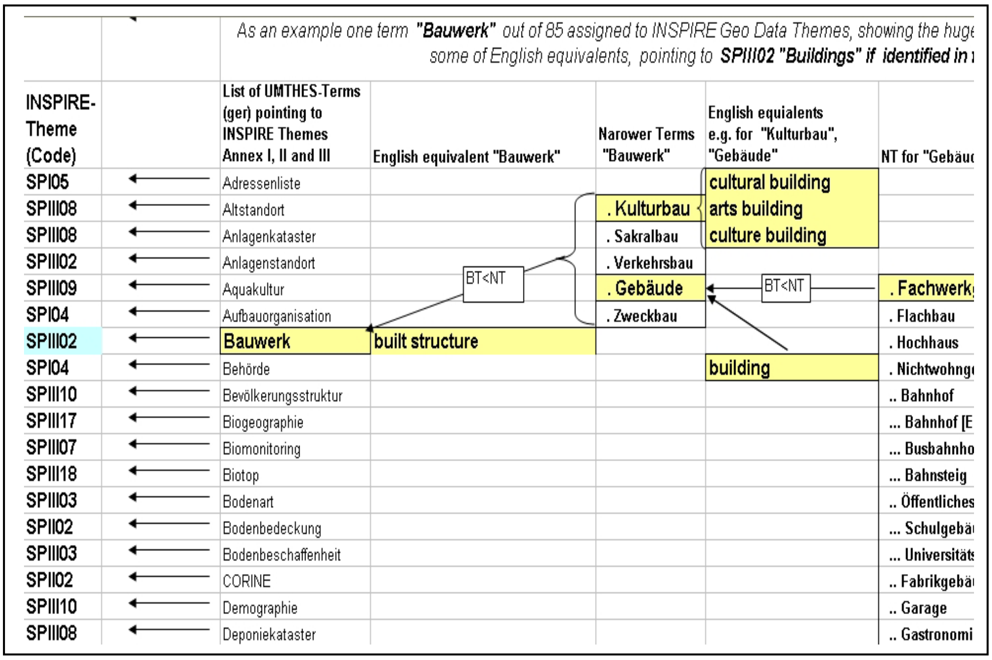


Figure 3: Example for assignment of UMTHES-Terms to INSPIRE Themes

With this step we may hope to make the meta data generation for INPIRE-Data more comfortable in this point.

REFERENCES

COMMISSION REGULATION (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata = VERORDNUNG (EG) Nr. 1205/2008 DER KOMMISSION vom 3. Dezember 2008 zur Durchführung der Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates hinsichtlich Metadaten

DIRECTIVE 2003/4/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 28 January 2003 on public access to environmental information …

COMMISSION REGULATION (EC) No 1205/2008 … of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata = VERORDNUNG (EG) Nr. 1205/2008 DER KOMMISSION vom 3. Dezember 2008 zur Durchführung der Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates hinsichtlich Metadaten

EEA 2008: Annual report 2007 and Environmental statement 2008 (Corporate document No 1/2008)

INSPIRE Directive 2007/2/EC - … OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)

Sharing environmental information through multilingual terminological and multimedia resources: the role of accessibility in increasing public awareness towards sustainable growth

Clara I. López-Rodríguez[a], **Juan A. Prieto-Velasco**[a], **Maribel Tercedor**[a]

[a] *Department of Translation and Interpreting, University of Granada, Buensuceso, 11, E-18071 Granada, Spain (clarailr@ugr.es, japrieto@ugr.es, itercedo@ugr.es)*

Abstract: One of the main challenges of environmental policies is increasing citizens' awareness and commitment towards our planet through effective communication. Effective communication between Member States and the general public can be achieved by reorganizing the sea of environmental resources available on the web in a systematic, accessible, and user-friendly way. However, some institutional websites in Europe tend to be programmatic, regulatory and not always appealing to citizens and their concerns.

The aim of this paper is to propose accessibility guidelines for the principled presentation and dissemination of environmental information in institutional websites. These guidelines focus mainly on three aspects: (a) accessibility both for users with different cognitive and physical abilities [Tercedor et al. 2007], and for users with different social, cultural or technological backgrounds; (b) increasing the amount of multilingual, multimodal and terminographical information so that non-experts can understand texts from governmental and non-governmental sources; and (c) translating and editing existing environmental campaigns and educational tools, which unfortunately are published in only one language, thus limiting access to potential users [Prieto et al. 2008].

In conclusion, to achieve a single space for environmental information in Europe, institutional websites should include accessible information that can be understood not just by experts, but also by the general public, should meet final users' expectations and should increase public awareness towards sustainable growth.

Keywords: Dissemination of environmental information; accessibility in institutional websites.

1. INTRODUCTION

Since 2005, the i2010 strategy is the EU policy framework for the information society and media. It promotes the positive contribution that information and communication technologies (ICT) can make to the economy, society and personal quality of life. This strategy has three aims: (a) to create a Single European Information Space; (b) to strengthen investment and innovation in ICT research; and (c) to support inclusion, better public services and quality of life through the use of ICT. Certainly, ICT can contribute to bridge the gap (i) between advances in Environmental Science and decision making processes and (ii) between scientists/policymakers and the expectations and interests of citizens.

The Internet plays an important role in achieving these objectives. Although the Internet has been defined as "the People's medium", not everyone has easy access or can interact with the Internet. Accessibility is then key to use the Internet and to promote inclusion in our information society. Moreover, taking into account the efforts made towards achieving a Single Information Space for the Environment in Europe, it is necessary to promote best practices in the sharing of environmental multilingual information taking advantage of online open access information sources.

Accessibility can be defined in a narrow sense or in a broad sense. The narrow sense implies that accessible content can be used by someone with a disability. However, this definition has been extended to include users with different cognitive and physical abilities, and social, cultural and technological backgrounds (users with minimum computer literacy, with different web browsers or channels to access the Internet. In fact, the World Wide Web Consortium [2003] defines accessibility as "the art of ensuring that, to as large an extent as possible, facilities (such as, for example, Web access) are available to people whether or not they have impairments of one sort or another". The W3C (Ibid) illustrates this point mentioning Internet users with differing needs and limitations, for instance: (i) users that may not be able to see, hear, move, or to process some types of information easily or at all; (ii) users having difficulty in reading or comprehending text; (iii) users that may not have or be able to use a keyboard or mouse; and (iv) users that may have a text-only display, or a small screen. We could add to these examples, users with limited knowledge of the language of the website, people who access the Internet through mobile devices (PDA, cell phones, Terrestrial Digital Television, etc.). All in all, accessible content is content that is easy to use by different people regardless of disability.

The W3C supports the idea of universal accessibility, and through its Web Accessibility Initiative (WAI) has published the Web Content Accessibility Guidelines (WCAG). These guidelines comprise a set of Checkpoints ranked into three categories, defined by the WAI as Priorities 1, 2 or 3, according to their relative decreasing importance in enabling Web access by people with impairments. Sites with Priority 1 receive the designation "Conformance Level A", those with Priority 2, "Conformance Level AA", and the maximum level of accessibility is given to websites with "Conformance level AAA".

In any case, accessibility benefits everyone regardless of ability. According to a survey carried out by the Disability Right Commission of Britain in 2004, accessible sites are 34% quicker to navigate. In their survey, a representative sample of websites used by British people were analysed according to their compliance to the WCAG 1.0. Altogether 1,000 webs were analysed with software that tests compliance automatically. This automatic analysis was enriched by a qualitative analysis of 10% of these sites carried out by users with different disabilities and accessibility experts. The survey concluded that 81% of the pages failed minimum standards set by the 3WC. The human evaluators of the websites reported the most frequent violations to WCAG accessibility checkpoints (Table 1). They also reported confusing and disorienting navigation mechanisms, cluttered and complex page structures, graphics and text size too small, and complicated language or terminology.

Table 1. Violations to accessibility in websites (Disability Right Commission 2004)

Provide a text equivalent for every non-text element [Priority 1]
Ensure that foreground and background colour combinations provide sufficient contrast when viewed by someone having colour deficits or when viewed on a black and white screen [Priority 2/3]
Ensure that pages are usable when scripts, applets, or other programmatic objects are turned off or not supported. If this is not possible, provide equivalent information on an alternative accessible page [Pr. 1]
Until user agents allow users to freeze moving content, avoid movement in pages [Priority 2]
Until user agents allow users to turn off spawned windows, do not cause pop-ups or other windows to appear and do not change the current window without informing the user [Priority 2]
Divide large blocks of information into more manageable groups where natural and appropriate [Pr. 2]
Clearly identify the target of each link [Priority 2]
Use the clearest and simplest language appropriate for a site's content [Priority 1]

Thus, if we want the i2010 strategy to succeed, policymakers and website developers should bear accessibility in mind, and involve both general users and disabled users from an early stage in the design process. Although there are new tools[1] that test accessibility automatically and are useful to correct violations of the WCAG, they do not detect all the accessibility problems. Therefore, the qualitative and quantitative opinion of users is of prime importance and should be taken into account by governments and environmental websites.

However, involving users in the design of accessible sites is not an easy task. Considering that the two versions of the WCAG are complex for the uninitiated and do not always follow a logical structure, it is necessary to draft a user-oriented checklist that highlights the accessibility deficiencies and communicative barriers of European environmental websites. Consequently, the aim of this paper is to propose a checklist that can be used to test to what extent institutional web pages actually reach citizens and increase their awareness towards sustainable growth. Our guidelines include a list of easy to understand points based on the guidelines and best practices of the W3C (How to Meet WCAG 2.0, http://www.w3.org/WAI/WCAG20/quickref/), the BBC (Future Media Standards & Guidelines: Accessibility, http://www.bbc.co.uk/guidelines/futuremedia/) and the Equality and Human Rights Commission of Britain (http://www.equalityhumanrights.com).

2. USER-ORIENTED GUIDELINES FOR ACCESSIBILITY

We have elaborated a questionnaire with 5 sections and 65 closed questions aimed at Internet users regardless of their computer literacy. The questionnaire will be used to test the communicative efficiency and accessibility of environmental web pages. The possible answers are: Yes (Y), No (N), or Not sure (NS). We understand that a positive answer to more that 50% of these items (simple majority) would be considered a good score in the evaluation of websites as far as accessibility is concerned.

Section 1 describes how a web page is accessed and how first impressions influence the decision to remain in the webpage. Section 2 focuses on navigation and operability, in other words, it focuses on whether the web provides ways to help users navigate, find content, and know where they are. Section 3 deals with perception of the web page through different sensory channels. In section 4, we concentrate on how the content on the web page is understood, and the role of language in making a web page comprehensible. Finally, Section 5 accounts for the interaction between users and the web, and analyses compatibility of formats and assistive technologies.

Table 2. User-oriented guidelines for accessibility

1. Access to the Homepage and first impressions	
a. The web was easily found with a search engine.	
b. The web embeds accurate metadata (HTML tags with information about the page) that facilitate the work of search engines, in particular, it provides title for the webpage that is informative, precise, and readable by users. The title appears in the title bar of a browser.	
c. The web embeds accurate metadata (HTML tags with information about the page) that facilitate the work of search engines, in particular, keywords metadata (up to 50 keywords) and description metadata about the content of the web.	
d. The web embeds accurate metadata (HTML tags with information about the page) that facilitate the work of search engines, in particular, description metadata about the content of the web.	
e. The web is attractive to users.	
f. The web provides double access for experts and the general public.	
g. At first sight, the page seems well organised.	
h. The page fits the screen without scrolling down.	
i. The web includes an overt mention to its conformance level of Accessibility (W3C-WAI).	
j. The web includes a CSS (Cascading Style Sheet) that facilitates the interaction of assistive technology with content by logically separating the *content's structural encoding* (indication of elements such as headings, paragraphs, lists, tables, etc.) from *presentational encoding* (formatting effects, such as typeface, colour, size, position, borders, etc.).	
2. Navigation and operability	
a. The web shows properly on different web browsers (Text-only, Mozilla, Internet Explorer, etc.).	
b. It is possible to navigate the web with the keyboard. The web provides access keys that can be used as shortcuts.	
c. Skip to navigation: for people using screen readers the first link on every page says “skip and go to site navigation”. Following this link will skip over the page header and go to the main site navigation menu. The page header is repeated on every page so this avoids hearing it again and again.	
d. Skip to content: for people using screen readers the second link on every page says “skip to content”. Following this link will skip over the page header and skip over the navigation menu to take you straight to the content of the page.	
e. Information about the user’s location within a set of Web pages is available in the form of a site map of the main headings of the website, an A-Z Index of topics covered in the web site or a clear hierarchical organisation.	
f. Searching for specific information: users can search the web site by typing one or more keywords into a search box.	

g. Advanced search options: if the search box retrieves too many results, there is an advanced search option to refine your search.	
h. Search tips: there is further help for using search in the Search tips.	
i. Intelligent searching: there is an intelligent search engine that retrieves semantically-related data because it is based on ontologies.	
j. Specification of Link Purpose: the target or purpose of each link is clearly identified from link text alone.	
3. Perceivability of the web page	
a. The web provides text alternatives for any non-text content so that it can be changed into other forms people need such as large print, Braille, speech, etc.	
b. The web provides an alternative text for any non-text content (image, image map, etc.) by using the ALT attribute.	
c. The web provides alternatives for any non-text content by using the LONGDESC attribute. This attribute specifies a link to a long description of an image (painting, chart, graph, etc.). This description should supplement the short description (maximum of 80 characters) provided using the ALT attribute.	
d. The ALT and/or LONGDESC attributes are used appropriately in order to describe visual content. When visual elements are merely decorative, these attributes are empty and can be ignored by assistive technology.	
e. The web provides alternatives for time-based media, in particular, subtitles for videos and animations.	
f. The web provides alternatives for time-based media, in particular, audio files for videos and animations.	
g. The web provides alternatives for time-based media, in particular, audio description for videos and animations.	
h. The web provides a sign language version.	
i. It is possible to change the size of the text.	
j. It is possible to change the colour of the website so that foreground and background colour combinations provide sufficient contrast for people with colour deficits or people viewing the page on a black and white screen.	
k. It is possible to freeze moving or sound content, or to avoid movement in pages.	
l. It is possible to turn off spawned windows or to prevent pop-ups or other windows to appear.	
m. You can hear the web page read aloud because the web implements programs such as Read Speak, Browse Aloud, Dixerit or Loquendo.	
n. The frames of the web are clearly delimited, and allow a natural reading process (firstly, you read the top of the page, then the left, and finally the centre).	
o. The different pages have a similar design and provide a sense of coherence and unity.	
p. The web makes it easier for users to see and hear content by separating foreground and background.	
q. The web does not include sensory features that are known to cause seizures.	
4. Understandability of the webpage	
4.1 Content	
a. The quantity and depth of information is appropriate for the site's intended user.	
b. Quality of information: there is not duplicity of information.	
c. The web can only be understood by experts.	
d. The web contains resources for children.	
e. The web contains audio files.	
f. The web contains videos/animations.	
g. The content of the web page is coherent.	
h. The web provides users enough time to read and use content.	
i. The web is not crammed with logos.	
4.2 Linguistic Aspects (grammar, style and multilingual aspects)	
j. The page displays good writing style presenting information in a plain and simple language, appropriate for the site's content. Administrative jargon is avoided.	
k. The web does not contain spelling mistakes.	
l. The web includes multilingual contents.	
m. Appropriate and fluent translations whenever they appear.	
n. The text has a clear and logical structure. Paragraphs are well constructed: one paragraph equals one idea.	
o. There is a glossary that explains difficult words and phrases.	
p. In the glossary, the language of the definitions is simple and easy to understand by non-experts.	
5. Interaction and compatibility with user agents and assistive Technologies	
a. The web includes well-developed formularies (brief, understandable) with hints to fill them out.	
b. Some documents are available in a variety of formats.	
c. It is possible to download contents and documents of the webpage for later offline use.	
d. It is possible to order media and documents for delivery in accessible format.	
e. There is help with accessing PDF files.	
f. There is a clear section of News and updates.	
g. The web provides a "Syndicate this site" system.	
h. Pages are usable when scripts, applets, or other programmatic objects are turned off or not supported.	
i. Uploading and downloading files does not take long.	
j. It is possible and easy to contact a helpline or the Webmaster by writing an email via the website.	
k. The page includes interesting and useful links.	
l. The page contains useful links to accessibility web sites.	

In this checklist, special attention has been paid to sections 3 and 4. We believe that some flaws of institutional websites are scarce accessible alternatives for people with disabilities, lack of multilingual contents and lack of didactic aids for non-experts in the form of glossaries, conceptual trees, multimedia materials, etc. We also believe that these guidelines can be a starting point towards the design of appealing, really accessible and understandable websites about the environment.

3. ACCESSIBILITY IN INSTITUTIONAL SITES: LEGISLATION AND ITS APPLICATION IN ENVIRONMENTAL ADMINISTRATIONS

Nowadays web accessibility is a legal requirement in the European Union. European countries have generally ratified the United Nations Convention on the Rights of Persons with Disabilities (2006), and the Council Resolution on Accessibility of Public Websites and their Content (2002), according to which "persons with disabilities include those who have long-term physical, mental, intellectual or sensory impairments which in interaction with various barriers may hinder their full and effective participation in society on an equal basis with others". These pieces of international law are legally binding for any country that has ratified them such as Germany, Spain, and the United Kingdom. These three countries have implemented their own national laws on accessibility, which are primarily founded on the W3C Web Content Accessibility Guidelines:

- Germany: Law on Equal Opportunities for Disabled Persons *(Gesetz zur Gleichstellung behinderter Menschen und zur Änderung anderer Gesetze 2002)* and the Ordinance on Barrier-Free Information Technology *(Barrierefreie Informationstechnik-Verordnung)* by the Federal Ministry of Labour and Social Affairs.
- Spain: Law on Equal Opportunities, Non-Discrimination and Universal Accessibility for Persons with Disabilities *(Ley de Igualdad de oportunidades, no discriminación y accesibilidad universal de las personas con discapacidad, 2003)*.
- United Kingdom: *Disability Discrimination Act* (1995; 2005).

In compliance with these rules, the national governments of Germany, Spain, and the United Kingdom have committed themselves to make accessible all institutional websites. People with disabilities on their part have the right to take legal actions against discrimination by websites and other information technology of domestic administrations not compliant to national accessibility laws.

In this section we apply our checklist to the analysis of the websites of the Departments for the Environment of Germany, Spain, and the United Kingdom, and see whether accessibility is a real concern. Accessibility should include navigation, perceivability, understandability, usability and interaction between the Web and people with different abilities and backgrounds.

3.1 Bundesumweltministerium (BMU)

The website of the German Ministry for the Environment, Nature Conservation and Nuclear Safety (*Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit*, http://www.bmu.de) can be considered accessible according to the Guidelines for Accessibility provided in Section 2. The results of the questionnaire show that the website fulfils 39 items out of 75; it includes 52% of features considered important for accessible institutional websites.

The first impression regarding access to the homepage is that of a properly structured and easy to understand website. It contains metadata useful for search engines and is built upon a CCS style sheet. Contents are pertinent and coherently presented. From a linguistic point of view, it uses a plain understandable language with appropriate style. Regarding the interaction with users, the website facilitates document download by presenting contents in several formats, and offers news and information in RSS (Really Simple Syndication standard for web feed) and significant links.

However, if we analyse the website in depth, we observe that the weakest aspect is that of perceivability, one of the most important in relation to multimedia contents, since the website does not include any subtitles, audio descriptions, textual transcriptions or sign language alternatives for videos, animations, or audio. Navigation meets only some of the recommendations included in our guidelines, so that users with disabilities may encounter several difficulties when browsing the website (lack of colour contrast, text-only browsing, lack of keyboard navigation, etc.).

To sum up, although apparently accessible, this website fails to meet core requirements for web accessibility: navigation may result difficult for people using screen readers and little help is provided for people with different types of sensory disabilities preventing them from easily accessing non-text contents.

3.2 Ministerio de Medio Ambiente y Medio Rural y Marino (Marm)

The website of the Spanish Ministry for the Environment (http://www.marm.es) is easily found by search engines because it includes accurate metadata. They specify the title of the page, its keywords and a brief description: "Información sobre la estructura gestión y política del Ministerio de Medio Ambiente y Medio Rural y Marino". The design of the page is attractive though not very original (green leaves on a blue sky) and consistent throughout the website. The homepage does not fit the screen without scrolling down, due to the fact that it includes too much information, photographs and logos (even of governmental campaigns not related to the environment). On the bright sight, the web contains a Cascading Style Sheet that facilitates interaction with assistive technology and claims a Conformance level of AA. This conformance level is not validated by online accessibility tests such as TAW. In fact, the web includes 57 breaches to Conformance level A and 89 to Conformance level AA. There is a significant difference between the design and accessibility features of the Homepage and those of the other pages.

As far as navigation is concerned, the Homepage scores low (3 out of 10), although when you click on the different sections, you find some navigation features not present in the homepage: search box, A-Z index of topics, etc. One of the main obstacles to navigation is the fact that some links share the same alternative text but they lead to different resources.

The perceivability of the web page should also be improved. The Homepage provides text alternatives for images (only through the ALT attribute; the LONGDESC attribute is inexistent), but in many cases it is not appropriate or precise. For example, the ALT attribute "Últimas noticias del gabinete de prensa del Marm" appears on 4 images (3 with a decorative function) tackling completely different topics. The alternatives for time-based media are restricted to subtitles for videos and animations, and are not used in a systematic way. In fact, the videos cannot be opened in a new window, thus minimising visibility. Fortunately, whenever a PDF file pops up, the web alerts the user about this. It is not possible to change the size or colour of the text. The frames and design of the web are not coherent throughout the site, and this may cause confusion.

As far as understandability is concerned, the content of the page is a bit chaotic, possibly because it provides too much information. From a linguistic point of view, the Homepage uses simple language indicating that it is aimed at the general public. However, the sections of the page mainly contain administrative jargon, legislation and propaganda. No glossaries or definition for non-experts are included. Regarding the interaction with users, the website facilitates document download by presenting contents in several formats (videos and PDF files) and offers news and information in RSS and numerous links (maybe too many).

3.3 Department for Environment, Food and Rural Affairs

The British Department for Environment, Food and Rural affairs (http://www.defra.gov.uk) has a site with a well structured menu bar with descriptions in tags. For instance, the menu *Plants and seeds* provides the descriptive content "plant health, varieties and seeds and issues relating to bees and honey production". Other alternative texts include ALT text for images pointing at the function or content of the image. For example, the institutional logo appears in all pages and, since it can be used as a pointer, it has the alternative text "Link to home page". When an image serves only decorative purposes, no ALT text is added. However, there are instances where images have a warning or preventive purpose, and provide no alternative text whatsoever, such as the poster explaining what food may be brought into the UK (http://www.defra.gov.uk/animalh/illegali/allow/whatfood.htm). By the same token, there is a video of a campaign to avoid bringing food into the UK (Don't bring me back campaign) and such video provides karaoke style captions of the song aimed at children, but no transcript of the speaker's voice warning of the criminal effects of personal food imports from outside of the EU.

The site is clearly aimed at the general public, as revealed by the language and style used. For instance, within the *About us* section, the headings "who are we?" "What do we do?", "How do we work?" indicate a language aimed at reaching the citizenship. All pages describing specialised topics such as animal illnesses first describe the condition to the general public in one paragraph. Even the sections containing legislation, usually begin by introducing the issue to the general public, and then provide links to the legislation itself. Some links may be confusing or ambiguous. For example, there is a *Media centre* with links to speeches but the contents are transcriptions, not actual video or audio files. A positive aspect in the *Contact us* link is that it provides contact information both from within the UK and from outside. When a link takes you to a PDF document, the text "opens a new window" is used in the link's descriptive text.

Moreover, there is a search bar with an alphabetical index and a glossary of abbreviations and acronyms. Such glossary contains only the complete form of abbreviations, but no definition or explanation is added, which would be helpful.

Regarding format, the homepage fits on the screen and there's no scrolling bar. The colour contrasts work well and make it attractive yet easy to read. Breadcrumbs such as "You are here: Homepage > Environmental protection" help you know where you are all the time. Finally, the accessibility section only contains information about key access for people with disabilities, following the British convention.

4. SHARING ENVIRONMENTAL INFORMATION THROUGH MULTILINGUAL AND MULTIMEDIA RESOURCES

Accessibility to specialised language resources arises from the fact that not all users can access specialized knowledge or easily interact with multimodal information, especially with multimedia documents, either due to legal, economic, technical or methodological barriers [Budin and Melby, 2000]. For these authors, interoperability is *conditio sine qua non* in order to increase knowledge sharing; interoperability implies a common and clear language with shared meanings, and the application of quality management standards and terminology management standards.

The analysis of institutional web sites in Europe clearly shows the need to enrich environmental information by achieving accessibility for users with different levels of knowledge, abilities and backgrounds [Tercedor et al., 2007], and the need to present environmental issues in a clear, simple and attractive way. In this paper, we have acknowledged the great potential of language in meeting accessibility targets. From this perspective, we believe that it is necessary to include multilingual information and terminographical resources in institutional web sites. This material is useful both for experts looking for the translation of a term into other European languages and for non-experts who want to understand specialised information. Following Faber et al. [2007], Faber and Reimerink [this volume] and León and Magaña [this volume], we claim that terminographical resources should be multilingual and multimodal, and should include (a) linguistic information (definitions, synonyms, equivalents in other languages, syntactic and collocational information); (b) conceptual information in the form of conceptual trees; (c) contextual information (concordances and contexts), and (d) multimedia resources (URLs, images, videos, animations, audio files, etc.).

Finally, considering that many interesting environmental campaigns and educational tools are published in only one language, European governments and institutions should seek the collaboration of European Faculties of Translation. These educational institutions could translate and edit environmental textual material, and provide glossaries with easy to understand definitions for the general public under the supervision of professional translators and experts from the environmental institutions involved.

5. CONCLUSIONS

To achieve a single space for environmental information in Europe, institutional websites should include accessible information that can be understood not just by experts, but also by the general public. In this paper, we have proposed user-oriented accessibility guidelines for the principled presentation and dissemination of environmental data in institutional websites about the environment. We have applied those guidelines to the description of the Departments for the Environment of Germany, Spain and the United Kingdom.

Accessible resources should guarantee the integration of a heterogeneous audience into the information society by providing equal opportunities to those with disabilities, poor reading abilities, less developed cognitive competence on the subject field or different social and technological backgrounds.

ACKNOWLEDGEMENTS

The authors wish to thank the Andalusian Regional Government for the financial support of the project *MarcoCosta: multilingual knowledge frames in Integrated Coastal Management* (P06-HUM-01489), and the Spanish Ministry for Science and Innovation for supporting the project *ECOSISTEMA: Single Information Space for Frame-based Environmental Data and Thesaurus* (FFI2008-06080-C03-01/FILO).

REFERENCES

Budin, G., and A. Melby, Accessibility of multilingual terminological resources: current problems and prospects for the future'. In Proceedings of the 2nd International Conference on Language Resources & Evaluation. Athens, Greece, 2000.

Disability Right Commission, The web: Access and inclusion for disabled people. DRC Formal Investigation report, 2004. Available at: http://www.equalityhumanrights.com/Documents/Disability/Accessibility_guidance/web_access_and_inclusion.pdf.

European Commission, The i2010 strategy. Available at: http://ec.europa.eu/information_society/eeurope/i2010/strategy/index_en.htm.

Faber, P., P. León, J.A. Prieto, and A. Reimerink, Linking images and words: the description of specialized concepts, *International Journal of Lexicography* 20 (1), 39–65, 2007.

Prieto Velasco, J.A., M.I. Tercedor Sánchez, and C.I. López Rodríguez, Using multimedia materials in the teaching of scientific and technical translation. *Linguistica Antverpiensia* 6. Special Issue on *Audiovisual Translation: A Tool for Social Integration,* 2008.

Tercedor Sánchez, M. I., C.I. López Rodríguez, and J.A. Prieto Velasco, Accesibilidad web y material multimedia: la descripción de imágenes. In B. Ruiz Mezcua and Francisco Utray (eds.). *Accesibilidad a los medios audiovisuales para personas con discapacidad*. Madrid: Real Patronato sobre Discapacidad, pp. 75 – 84, 2007.

World Wide Web Consortium (W3C), Web Accessibility Initiative Glossary: Internal Working Draft. Available at: http://www.w3.org/WAI/GL/Glossary/printable.html, 2003.

World Wide Web Consortium (W3C), Web Content Accessibility Guidelines 2.0. Available at: http://www.w3.org/TR/WCAG20/, 2008.

[1] These tools can test accessibility both off-line (Fujitsu Web Accessibility Inspector, http://www.fujitsu.com/global/accessibility/assistance/wi) and online (W3C Markup Validation Service, http://validator.w3.org; CTIC Test de Accesibilidad Web, TAW, http://www.tawdis.net/taw3/cms/es).

The Environment Ontology: Linking Environmental Data.
(Position Paper)

N. Morrison[a], **M. Ashburner**[b], **D. Field**[c], **S. Lewis**[d], **C. Mungall**[d], **L. Schriml**[e], **B. Smith**[f] **and The Environment Ontology Consortium**[g]

[a] *NERC Environmental Bioinformatics Centre, NERC Centre for Ecology & Hydrology, Oxford, UK and School of Computer Science, The University of Manchester, UK (norman.morrison@manchester.ac.uk)*

[b] *Department of Genetics, University of Cambridge, UK (ma11@gen.cam.ac.uk)*

[c] *NERC Centre for Ecology & Hydrology, Oxford, UK (dfield@ceh.ac.uk)*

[d] *Lawrence Berkeley National Laboratory, Berkeley, CA, USA (suzi@berkeleybop.org; cjm@berkeleybop.org)*

[e] *Institute for Genome Sciences, University of Maryland, School of Medicine, Maryland, USA (lschriml@som.umaryland.edu)*

[f] *New York State Center of Excellence in Bioinformatics and Life Sciences, University at Buffalo, USA (phismith@buffalo.edu)*

[g] *www.environmentontology.org*

Abstract: Scientific research is benefiting from an unprecedented explosion of data. There is an urgent need to systematically describe 'where' this data comes from. We present here the aims of The Environment Ontology Consortium; to facilitate the semantic interoperability of environmental information associated with biological data of any organism or biological sample, through the development of standards, ontologies and associated reporting guidelines.

Keywords: Environment; Ontology; Controlled Vocabulary; Terminology.

1. INTRODUCTION

Every biological specimen that is collected or sampled – whether for a museum collection, for epidemiological studies, for population studies, ecological research, for research into evolution, biodiversity or sustainability, indeed any biological research – comes from a particular habitat where particular physical conditions prevail. Knowing the specific conditions and locale from which these biological samples are collected is essential for retrieval of data and for making comparisons [Editorial, 2008]. Indeed, such comparisons are central to gaining an understanding of the controlling variables involved in the maintenance of living things.

2. MOTIVATION

At present there is no accepted semantic standard for describing the environment from where biological samples are collected. This is a serious problem for anyone wishing to retrieve and compare environmental data. Grounding the capture of biological data within an ontological framework is proving to be increasingly useful in the fields of biological and clinical research and we believe that the development of an environment *ontology* (which differs from a terminology in that it is intended to support computational reasoning) can bring similar benefits.

3. COMMUNITY AND DEVELOPMENT

Under the umbrella of The Environment Ontology Consortium (www.environmentontology.org) work has begun to provide an integrated approach to the problem of linking environmental data. The aims of this effort are to support the semantically consistent description of, and computational reasoning over, environmental information associated with biological data for any organism or biological sample.

The task of describing the environments of organisms and biological samples has been divided into two orthogonal yet complementary sub-projects. EnvO is an ontology that describes environment types. Gaz represents a first step towards an open source gazetteer, constructed on ontological principles, that describes places and the relations between them. Combined, an environment ontology and associated gazetteer – in which place names are annotated with environmental information and with GPS coordinates – will provide a format that can be read and used by software agents, thus permitting them to find, share and integrate information that ordinarily would have required human intervention.

Currently EnvO and GAZ are represented using the OBO Language, which has a corresponding representation in the Web Ontology Language (OWL). We chose this representation over thesaurus standards such as SKOS as we intend to eventually support computational reasoning, using relationships encoded within EnvO, and connecting to other ontologies within the Open Bio Ontologies (OBO) Foundry (http://www.obofoundry.org) [Smith *et al.*, 2007].

4. APPLICATION

The application of EnvO and Gaz can serve as a test bed for establishing robust guidelines for biologists and others recording information about environmental phenotype and other geo-spatially indexed locations. Potential domain applications include epidemiological studies, studies of immigration and emigration patterns, studies of aspects of human environments related to health and disease, studies of the effects of climate change on agriculture and livestock; all of which require controlled vocabularies for describing environmental information in association with spatial descriptions of biological phenomena of different sorts. These examples highlight just some of the areas in which an Environment Ontology would be of benefit to the scientific community.

5. CONCLUSIONS

It is our hope that the recommendations of this community will build towards an integrated framework that can be used for the annotation and intelligent retrieval of any record that has an environmental component. We are an open access community and encourage stakeholders with similar interests to become involved.

ACKNOWLEDGEMENTS

The authors wish to thank all members of the Environment Ontology Consortium for their continued support in the development of a robust semantic framework for representing environments.

REFERENCES

Editorial, A place for everything, *Nature*, 453, 2, 2008.

Smith B, Ashburner M, Rosse C, *et al*. The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration, *Nature Biotechnology*, 25 (11): 1251-1255, 2007.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mirovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

A Challenge for the Global Water Management: The Harmonization of Terms, Definitions and Concepts

Michael Nagy

Unit Data Management & Reporting , Umweltbundesamt Wien, Austria,
michael.nagy@umweltbundesamt.at

Abstract: Harmonization of concepts, terms and definitions is essential for the development of a sustainable water management on local, national, European and the global level. There is a big number of stakeholders involved in water policies who have developed their own terminologies. This leads to incomparable indicators, inconsistent data sets and double-work carried out by different institutions.

In order to address this problem, international agencies as Eurostat or the United Nations have started an initiative to harmonize concepts, terms an definitions concerning water statistics and water data.

Keywords: Water statistics; Water Accounting; WISE; Water data

1. INTRODUCTION

Water is life! Therefore, protection of the water resources and the ensuring of sustainable use of water are among the first priorities on the agendas of international organizations such as the United Nations, the European Environment Agency or the European Commission.

To define goals and measures on a multi-country level a common “language”, including agreed terms and definitions, is needed. In the past water management was a national task, the individual countries and even different stakeholders within countries have developed different terms and definitions which are not always comparable between countries or regions. The indicator-based work of the European Environment Agency, the European Water Framework Directive (2000/60/EC) as well as the Eurostat/OECD-Joint Questionnaire on Inland Waters have supported very much the harmonization of terms and definitions among European Countries. However, work is not yet completed and terms as for example “water supply” or “wastewater” still are being understood in different ways. Consequently that may lead to misinterpretation when using data for international comparisons or policy making on the European level.

2. THE INTERNATIONAL HARMONIZATION PROCESS

Recently, within the frame of the implementation of the Water Information System for Europe (WISE), work has started to share water data that has been collected in different contexts. This is important to provide a Europe-wide picture on sustainable water use and the availability of water from natural water resources. For this task it is necessary to use data which has been collected in different contexts (like e.g. European Wastewater Treatment Directive 91/271/EEC, the European Pollutants Release and Transfer Register (e-PRTR) or statistical data as collected by Eurostat). It is very important for this work not only to harmonize concepts, terms and definitions, but also streamlining reporting flows in terms of frequency and QA/QC procedures. Following the SEIS principle “information should be collected once, and shared with others for many purposes” requires therefore an extensive debate amongst European Institutions (Group of 4), the EU-Member States and different stakeholders in water management to develop a terminology that allows the providing of the big European picture on sustainable water uses and the development of available water resources.

Similar discussions on global level, led by the United Nations, have been started recently. In the year 2005 the Intersecretariat Working Group on Environment Statistics (IWG-ENV) Sub-Group Water Statistics met in Vienna to discuss the harmonization of data collections among UN-Institutions (e.g. UNSD, UNEP, FAO) and other international institutions (OECD, European Commission, Mekong Commission etc.). One of the conclusions of this meeting was that "regional and international organisations actively collecting water data need to work more closely with each other and with UN Water to improve data availability and harmonization and to eliminate duplication. Responsibilities need to be clearly defined, so that data needs to be sent only once to international organization." For more details see http://www.umweltbundesamt.at/iwg-env.

3. INTERNATIONAL RECOMMENDATIONS FOR WATER STATISTICS (IRWS)

Since the adoption of the System of Environmental-Economic Accounting for Water (SEEAW, UNSD 2007) by the Statistical Commission in 2007 as an international statistical standard the process of harmonization of concepts, terms and definitions used in water management and water statistics has become a global priority. To facilitate this process the United Nations Statistics Division (UNSD) is currently elaborating International Recommendations for Water Statistics (IRWS) which are under an internal review by know.

There are 2 chapters in the IRWS which are important for the future standardization of the statistical units and data items:

Chapter 3: Statistical Units and Classifications used in Water Statistics

Chapter 4: Data items

3.1 Chapter 3: Statistical Units and Classifications used in Water Statistics

Chapter 3 introduces the classification of inland water resources that form the statistical units of the environment for which water statistics are collected. It furthermore, defines the institutional units, which includes the description of enterprises, establishments and households as well as the concept of residence.

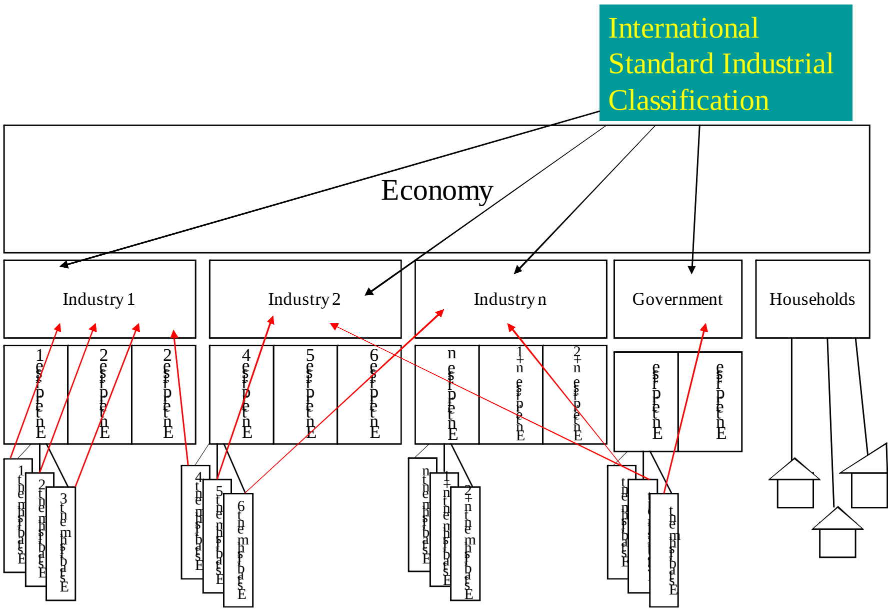
Example A: Statistical units of the environment:

- Surface water (EA 131)
    - Artificial reservoirs (EA 1311)
    - Lakes (EA 1312)
    - Rivers (EA 1313)
    - Snow, ice and glaciers (EA 1314)
- Groundwater (EA 132)
- Soil water (EA 133)

The above classification follows the SEEA asset classification. However, there are still some discussions among experts, e.g. whether artificial reservoirs are part of the environment or not. A re-classification would not only result in a break in existing time series on water uses but would also have consequences on the allocation of water uses to different kind of industries (the reason is that artificial reservoirs loose water due to evaporation and have a number of additional negative effects on the environment and certain economic sectors).

Example B: Relationship of enterprises, establishments and industry classification:

For water statistics and water management it is very important to allocate water uses, water losses and wastewater emissions to industries and households. IRWS recommends to use the so-called “International Standard Industry Classification” (ISIC) for this purpose. This is a well accepted standard classification in the economic and statistical society, but usually something new for people working in the water administration. This recommended standard leads to full consistency of physical water use data and related economic data, but there are also good practical reasons why other classifications used in water administration have their right to exist and should not be overruled. Ways must be found to translate between the different societies and to find algorithms which can be used to rearrange data according to different user needs. The following diagram shows the relationship of enterprises, establishments and industry classification (ISIC) as recommended in the IRWS:

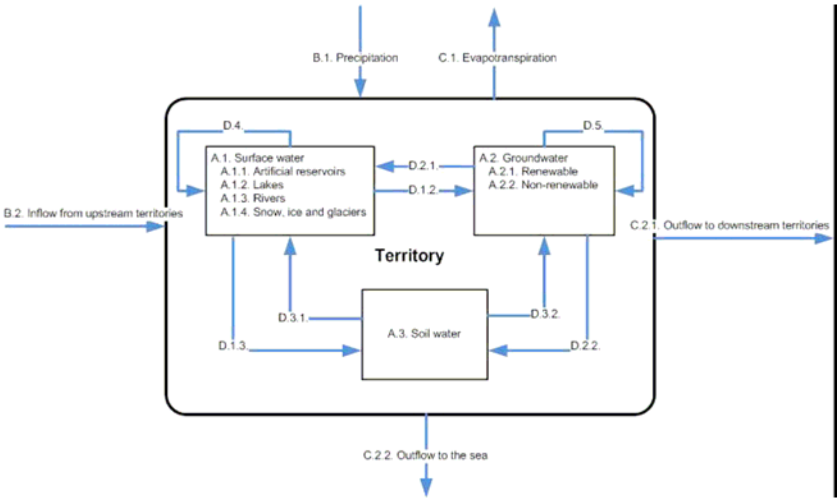


3.2 IRWS Chapter 4: Data items

Chapter 4 of the IRWS provides a comprehensive list of data items and their definitions recommended for collection, compilation and dissemination of water statistics and accounts. The data items are presented as part of a hierarchical classification and they constitute the basic building blocks of water statistics.

Example C: Flows of water in the environment:

This example shows the natural flows of water in the environment in form of a diagram and in form of the related, hierarchically ordered, data items.



- B. Inflow of water to a territory's inland water resources
    - B.1. Precipitation
    - B.2. Inflow of water from upstream territories
        - B.2.1. Inflow of water from upstream territories
            - B.2.1.1. Secured through treaties
            - B.2.1.2. Not secured through treaties
- C. Outflow of water from a territory's inland water resources
    - C.1. Evapotranspiration from inland water resources
        - C.1.1. Evaporation
        - C.1.2. Transpiration from plants
    - C.2. Outflow of water to downstream territories and the sea
        - C.2.1. To downstream territories
            - C.2.1.1. Reserved by treaties
            - C.2.1.2. Not reserved by treaties
        - C.2.2. To the sea

Example D: Physical data items for flows from the environment to the economy:

This example shows the different flows of water from the environment to the economy. For most parts of the hierarchical structure is already an informal agreement on international level. However, some parts of it (as e.g. soilwater and collection of precipitation) are not yet fully agreed upon.

- E. Abstraction of water
    - E.1. From inland water resources
        - E.1.1. From surface water
            - E.1.1.1. From artificial reservoirs
            - E.1.1.2. From lakes
            - E.1.1.3. From rivers

      - E.1.1.4. From snow, ice and glaciers
    - E.1.2. From groundwater
      - E.1.2.1. Renewable groundwater
      - E.1.2.2. Non-renewable groundwater
    - E.1.3. From soil water
  - E.2. Abstraction from other sources
    - E.2.1. Collection of precipitation
    - E.2.2. From the sea

Alternative breakdown

- E.a. For own use
- E.b. For distribution

The breakdown of water flows can be further disaggregated according to different water qualities (e.g. freshwater, brackish water).

4. CONCLUSION

The presentation will give an overview on latest international developments concerning the harmonization process of concepts and terminologies. There will be examples showing the possibilities to link existing concepts to each other in order to provide a comprehensive picture for global, regional and national water policies as well as for the presentation to different expert groups and the public. The presentation will put the IRWS and European activities into focus and show on the basis of examples where (informal) international agreements already exist and which parts of the international harmonization efforts are still unsolved.

REFERENCES

EUROSTAT (2007): Data Collection Manual for the OECD/Eurostat Joint Questionnaire on Inland Waters Tables 1 – 7. http://circa.europa.eu/Public/irc/dsis/envirmeet/library?l=/meeting_archives_1/meetings_2007_archive/statistics_09-10102007/eurostat_statistics/_EN_1.0_&a=d

UNSD - United Nations Statistics Division (2003): Integrated Environmental and Economic Accounting. http://unstats.un.org/unsd/envaccounting/seea2003.pdf

UNSD - United Nations Statistics Division (2007): System of Environmental-Economic Accounting for Water. http://unstats.un.org/unsd/envaccounting/SEEAWDraftManual.pdf

UNSD - United Nations Statistics Division (2008): DRAFT International Recommendations for Water Statistics. http://unstats.un.org/unsd/envaccounting/irws/egm2008/AC170-4.pdf

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

A state-of-the-art of Italian National Research Council (CNR) activities in the area of terminology and thesauri

P. Plini, **S. Di Franco, V. De Santis**

Italian National Research Council – Institute for Atmospheric Pollution - Environmental Knowledge Organisation Laboratory (plini@iia.cnr.it)

Abstract: In this paper the activities in the field of terminology and thesauri carried out by the Environmental Knowledge Organisation Laboratory of the Italian National Research Council will be presented. In 2000 EKOLab started the development of a thesaurus for the environment called EARTh (Environmental Applications Reference Thesaurus). The thesaurus is constantly updated following the growing request of updated terminology in relevant topics. At present it contains about 12.000 terms, in English and Italian. It is provided with hierarchical relations and classified according to a matrix system. In parallel with the development of EARTh, two separate projects were carried on and lead to the development of two thesauri, one on Remote Sensing and GIS and the other on the Snow/Ice environment. Each one of these thesauri contains about 3.500 terms. A 1.500 terms lexicon on pollution and climate changes has recently been created. The Thesauri are handled using SuperThes, a software developed within an International co-operation including CNR, UBA-Austria, UBA Germany and the TBHS company. The web interface is under development.

Keywords: Environmental terminology; Environmental thesaurus; Remote Sensing; Geographic Information System; Ice and snow environment; Conceptual model.

1. EARTh vs. GEMET

1.1 GEMET

GEMET, the GEneral Multilingual Environmental Thesaurus, was developed by CNR and the German Umweltbundesamt, between 1995 and 1999, as an indexing, retrieval and control tool for the Catalogue of Data Sources (CDS) of the European Environment Agency (EEA), Copenhagen. GEMET was conceived as a "general" thesaurus, aimed to define a common general language, a core of general terminology for the environment. GEMET was compiled by merging the terms of several multilingual documents. The merging has been performed both on conceptual and formal basis. Coinciding concepts in the different thesauri were identified and scored. From the point of view of translations, GEMET provides a complete equivalence (all the descriptors have an equivalent in a different language).

1.1.1 Criteria for the allocation of terms to the groups and themes of GEMET

GEMET has two systems for arranging the descriptors:

- A classification scheme of 3 super-groups containing 30 groups; there are in addition 5 accessory groups of terms, instrumental to the thesaurus use. The super-groups were adopted to approach an environmental management perspective and to help the hierarchical structuring of GEMET. The groups reflect a systematic, category- or discipline-oriented perspective. Within the groups, the descriptors were basically allocated in a mono-hierarchical order, but several descriptors needed to be allocated to more than one group or to more than a broader term inside the same group, thus creating a condition of poly-hierarchy.

- A thematic order, containing 40 themes. These themes were established according to practical considerations, corresponding to the information needs. They were developed to reflect the EEA activities in order to support the thematic elements of the EEA DPSIR Dataflow Scheme. The themes, being complementary to the groups, confer to the thesaurus a matrix structure.

Associative relationships were established between terms belonging to different logical categories. From the numerous possible relationships only those relationships which were considered useful for indexing and searching were included in GEMET.

In GEMET, the classification scheme by groups, themes and hierarchies, was intended merely as a mean to control the thesaurus terms and the semantic relations between them; in other words, only as a way to control the internal coherence of the thesaurus. As such, it was not proposed as a general reference pattern for the organisation of any specific environmental information system.

1.2 From GEMET to EARTh

1.2.1. Theoretical assumptions

Once concluded the work on GEMET, the EKOLab work in the field of knowledge organisation related to terminology and thesauri continued.

While representing the semantics of a term it is necessary to choose which types of characteristics and how many of them need to be considered and included in the representation.

In our approach it was considered that on one hand there is the necessity to share a common and stable meaning of the terms in order to guarantee communication within a community. Nevertheless, openness to a further exploration of meaning should also be ensured so as not to impoverish its richness and complexity.

In the current historical-cultural context of the Western tradition, there is one semantic trait that cannot be cancelled, unless we want to incur in a complete reformulation of the meaning of the term. Highly structured and refined, but flexible tools are in fact needed in order to deal with issues such as information management on the web or to satisfy the growing demand of semantic interoperability.

Our goal was the development of a thesaurus that could include the above assumptions aiming to become an advanced tool to be applied in environmental information management.

In order to sustain environmental policy and research, not only access to- but also high level quality of information is required. To achieve this result, however, systems capable of dealing with the specific features of the environmental sector are needed. The incompleteness of the terminology collections often depends, in fact, on the kind of approach that is utilised. Environment is mainly analysed with a static and sectorial approach, reflecting a vision pertaining to classic science and to environmental policy that transforms its paradigms in operational terms. This implies limited openings to the development of renewed approaches and methods to analyse the environmental issues. Moreover there is the constant need to follow the terminological changes in science and technology (new terms and/or new meanings, new topics and issues).

Starting from this premise, we adopted a more inclusive approach concerning both conceptual coverage and semantic organisation, and taking also into considerations suggestions arising from the development of applied ontologies, we started to work on an environmental thesaurus format that contains some innovative elements.

1.2.2. EARTh and its semantic model

EARTh (Environmental Application Reference Thesaurus) is based on a multidimensional classificatory and semantic model. The "vertical structure" of the Thesaurus is the fundamental constituent of such a model. This structure is basically mono-hierarchical. It has been developed according to a tree semantic model and is founded on a system of categories. It is organised in a framework composed of different levels and classification knots and comprises hierarchical relationships.

The notion "category" has taken on different meanings in the history of Western thinking; from the ontological and logical point of view to linguistics, from philosophy to semiotics to psychology. In the classification science categories have also been considered as the foundation (not always visible) of knowledge organisation systems and are utilised for different purposes [Barite, 2000].

In the context of the present work, categories are conceived in their primitive Aristotelian form as the most general *genera* or the logical progenitors under which every single terms can be placed.

The first two levels of the classification correspond to the system of categories. The first level includes four "super-categories": Entities, Attributes, Dynamic aspects and Dimensions. Entities constitutes "things". Attributes defines character of "things", at least in their static aspects. Dynamic aspects relates to transformations and operations connected to "things". Dimensions identifies the spatio-temporal circumstances where all this is manifested.

In the subsequent level of the classification, the super-category Entities is divided into Material entities and Immaterial entities. Attributes includes three different categories: Properties; Structure and Morphology; Composition. Dynamic aspects comprehends: Processes; Conditions; Activities. Dimensions refers to Space and Time.

According to this model the semantics of the terms is, in fact, described by the categories where they are located. Following a bottom-up perspective, terms could be analysed according to a progressive hierarchical scale. In that scale conceptual features are progressively discarded following an intensional perspective, while in an extensional perspective the number of things associated to that intension is increased. The maximum level of generality is thus reached. Categories represent the top of this vertical structure [Fugmann, 1993] that analyses the meaning of the terms according to a logical perspective. It can be considered as an operative tool that – by providing the categorial interpretation of the meaning of the terms and by placing them in the classificatory-hierarchical tree – aims to orientate the users towards the most "essential" characteristics of their semantics.

Nevertheless it does not limit the conceptual analysis of terms to a static and univocal view. Awareness of the semantic complexity associated with each term is maintained. Different layers of meaning have to be explored, even if there is a hierarchy of semantic traits and each one of them contributes to lexical signification with a different specific weight.

1.2.3. The Thematic setup

The model envisages the possibility to develop additional arrangements of the terminology according to the specific needs of the applicative context. A thematic organisation of terms has been elaborated. A theme or a subject is here conceived as a sector of interest that reassembles the terms related to it, while a tree or faceted structure tends to scatter them under their referral logical category.

In EARTh we have developed a thematic classification that has been utilised to classify the terms and that could also be used for the management of information in the field of research, environmental policy, dissemination to users.

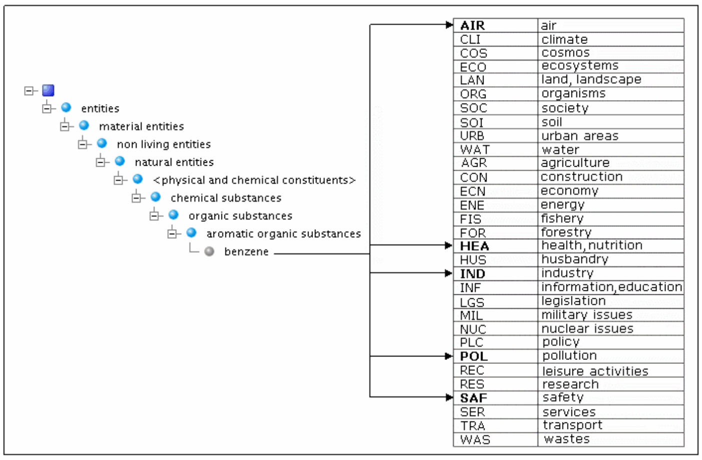


Figure 1. Classification of benzene according to hierarchical structure and to the systems of themes

From a semiotic point of view, this model should allow meaning representation according to different “second order” perspectives and acceptations. The possibility to apply additional classification models would ensure, in fact, openness and flexibility to the model. The RT relation can, of course, help in expressing additional semantic traits.

The topic of semantics of terms is more complex that it would appear. The proposed semantic model has been developed for operational purposes and is based on a regulating principle which feasibility needs to be further evaluated.

In the vertical structure, for example, the lower down the semantic hierarchy the more the tree becomes affected, at a general level, by historical and cultural contingencies and, more specifically, by operational choices and by the characteristics of the domain of knowledge.

2. EARTH RELATIONAL STRUCTURE

The usefulness of a well-structured domain-specific thesaurus for the management of information is rather acknowledged. However there is a widespread opinion that the traditional thesaurus format doesn’t completely fit the current needs. One of the main problems posed by thesauri seems to be the fact that they provide a poorly differentiated set of relationships between terms, distinguishing only among hierarchical relationships, associative relationships and equivalence relationships. It has been also said that since thesaurus relationships are characterised by semantic vagueness, they are not applied consistently. This causes ambiguity in the interpretation and can result in unpredictable semantic structures [Soergel, 2004]. Moreover thesauri are expected to be developed on the basis of a more fully concept-oriented model (while a term-oriented model, according to this viewpoint, may promote ambiguity and incompatibility) where concepts are considered to be independent and precede their designations.

In our project we tried to implement a more refined set of semantic relationships. Standard relationships will be arranged into richer subtypes, whose semantic content is specified. Linguistic structures will express semantic relations. The augmentation of thesaurus relationships will ensure a stronger semantic control, also because different relationships can hold each other in check [Fisher, 1998], and open up new possibilities for information retrieval applications [Tudhope, 2001]. The enrichment of the relations and their increased semantic clarification could enable, for example, a better semantic description of Web resources and guide a user in meaningful information discovery on the Web [Soergel, 2004]. Besides, it will increase the possibility of using them also for artificial intelligence applications. Traditional thesauri, in fact, were not designed for- and their semantic structure supports limitedly automated information processing [Soergel, 2004].

Nevertheless from another point of view we hold the awareness of the intrinsic complexity of contemporary lexicons that are systems very rich of polysemies, redundancies and so on. We have to provide tools able to ensure a stronger semantic control as much as possible. But we have also to use great care in avoiding an excess of “compulsory way” or artificially compressed meanings. While applying a highly elaborated net of semantic relationships unwanted effects of this kind could, in fact, be generated. It seems instead reasonable to adopt a hermeneutical attitude open to “accept”, to a certain extent the “weak” nature of lexicons.

Ensuring a high modularity of these systems is another important requisite to be achieved. This should enable also other kinds of utilisation by users that may not need or make such a fine distinction of the thesaurus relations [Milstead, 2003], in this case it could even become a problem more than a solution, and that are interested in using a simpler and more traditional version of the thesaurus relational structure.

2.1 Hierarchical Relationship

Thesaurus standards and the scientific literature include three kinds of hierarchical relations: “generic”, “partitive” and “instance”, which are conflated into one generic ”hierarchical relationship”. Perhaps this is the most misused relation**.** Many existing thesauri, claiming to be ISO standard consistent, provide relations that are labelled as BT/NT but they could be better interpreted as associative relations. They are, in fact, based on a document-retrieval definition of ‘broader-narrower’ that is of pragmatic nature and oriented towards the function of the search process [Fisher, 1998]. In EARTh only hierarchies that are logically based will be included. Moreover we will differentiate the different types of relations and as a second step subtypes will be identified.

2.2 Associative Relationship

The associative relation is quite difficult to describe because it covers a heterogeneous and undifferentiated set of relations. ISO 704 defines it as a relation that “exists when a thematic connection can be established between concepts by virtue of experience”. It can express many kinds of association between terms that are not hierarchically based. Such links should be made explicit in a thesaurus since they suggest additional terms that can be used in indexing or retrieval.

In our work we will try to specify the nature of the relations and to differentiate RTs in subtypes (i.e., “cause/effect”, “raw material/product”, “discipline/practitioner”, etc.). We will also try to extend the range of useful RTs types although they probably constitute a series that is intrinsically open and their making topical is strongly connected with the characteristics of the operative context.

In this way, by strengthening the transversal relational structure, which is based on associative relations, a knowledge representation model that is net-like structured is being developed. It will emphasize the system of interrelations, the “connecting” ties that limit the degree of separation of a conceptual field and cannot be represented by the taxonomic-hierarchic tree-like model [Trigari, 2003]. In our case this is very important also to obtain a system able to deal with the environment, which is a domain where the complexity of the systems as well as the web of interlinking, plays a key role. And it will also be useful to deal with the networked and barely hierarchical information and knowledge management on the Internet and to better reflect the emerging mental maps of the information searcher [Trigari, 2003].

To better represent and visualize this transversal structure, we are also thinking to the possibility of designing additional ways of browsing the thesaurus based on the RTs and showing different microworlds of connected concepts and terms [Trigari, 2003].

2.3 Equivalence Relationship

Equivalence relationship covers at least the following basic types: synonyms, lexical variants and near-synonymy. Synonymy refers to meaning similarity. It has also been defined as interchangeability between terms, although it is very difficult to think about the existence of an absolute or perfect synonymy where there is interchangeability in all contexts [Violi, 1997]. Classes of synonyms include, for example, dialectal variants, popular and technical term pairs, generic and trade name pairs, different linguistic origin variants, variant names for emergent concepts, slang or jargon synonyms and so on [Greenberg et al., 1997]. Lexical variants are different word forms for the same expression and derive from morphological and grammatical variations (i.e., orthographic and syntactic variants). For synonyms as well as for lexical variants we will try to identify different subtypes. The category of near-synonyms as such won't be included at this stage in the system.

3. MULTILINGUALISM AND CULTURAL DIVERSITY ISSUES

Semantic and structural divergences that may concern multilingual thesauri -here we are referring to regional diversities that occur in the context of a common general culture, as the Western culture- will be taken into account. Developing a multilingual thesaurus could require a redefinition of the relationship between the language versions included in the thesaurus. It could, in fact, lead to producing a non-symmetrically structured system, in which a full correspondence between semantic relations in different languages may not be achieved and the number of descriptors is not necessarily the same [Hudon, 1997]. According also to the Guidelines for Multilingual Thesauri [IFLA 2005] building a "symmetrical thesaurus" aims at a full correspondence between descriptors and relations. Not rarely, however, by adopting a symmetrical approach, cross-language equivalences are forced where they do not exist and questionable relational structures are established.

Moreover problems posed by cultures that are epistemologically far removed from the Western one, will also be considered. Nowadays connection at planetary level is strongly increased. Different cultures and knowledge forms meet on global platforms. We bring the awareness that different cultures hold different visions of the world and this is reflected in the way they organise knowledge too. Thus we are evaluating in which way the future development of the Thesaurus could best integrate a multicultural perspective.

4. EARTH POSSIBLE APPLICATIONS

EARTH was not originally intended as a thesaurus exclusively aiming at replacing the already existing environmental thesauri. Especially during the first phases of its development EKOLab invested relevant efforts in analysing the theoretical basis of thesaurus development and EARTh represented a suitable use case where to implement and test the research results.

Obviously this does not imply that EARTh can not be adopted and used as a reference thesaurus. Belonging to an institution whose primary task is the research, we are aware that EARTh is and will continue to be an "open" thesaurus where content updating and structure upgrading will continue to coexist.

4.1 Interoperability issues

The Thesaurus could be utilised for different purposes. It could be interesting to evaluate the use of EARTh, having in mind the structure of its semantic model, to deal with interoperability issues and as a tool for mapping among different environment-related thesauri. Nowadays networked information access to heterogeneous data sources requires interoperability of controlled vocabularies. Technology allows, in fact, shared access but incompatibilities in the vocabularies have to be solved so that they can interoperate for the purpose of information retrieval. Thesauri are created with different points of view, their development reflects different scopes and can be based on different ways of conceptualisation. They may differ in structure and in selection of concepts and terms.

Tools able to create dynamic and semantically based correspondences among different vocabularies are then urgently needed.

4.2 EARTh as environmental semantic map

Current research also indicates new roles of thesauri. They are increasingly seen as maps of subject domains or semantic networks, knowledge representation and organisation systems, patterns of knowledge [Kosovac, 1998]. Disciplines as applied linguistics, cognitive science and artificial intelligence can view and study them as models of conceptualisation and employed to structure knowledge domains according to logical and semantic criteria. A thesaurus that, like EARTh, is expected to have a refined semantic structure seems to be highly suitable for this scope and could allow domain-specific navigation on a semantic basis.

5. EARTH EXCHANGE FORMAT AND ACCESS BY SERVICES

At present EARTh database is handled using Firebird, an open source SQL database, in order to handle the thesaurus relations a specific software named "Superthes" was developed by TBHS, an Austrian company. At present data exchange may be performed in several ways:

- XML Export and Impor. SuperThes is also able to write and read this data exchange format. Benefits of this method is the general available document type definition and the possibility to read and write every detail of a Thesaurus. The drawback is the amount of work that must be put into a custom reader or writer.
- Import and export of "flat" (non-hierarchical) data directly to and from Microsoft applications. This method was already used with great success in porting GEMET into other languages. It may also be used to import any flat file (list of terms or similar) into the system.

There are two ways to make the thesaurus available to users over the internet, a web interface and a web service. The web interface was foreseen to be ready within the end of 2008, but actually it has been delayed and this features are expected to be implemented quite soon.

In the frame of a CNR Department of Earth and Environment project called GIIDA (Integrated and Interoperational Management of Environmental Data), a web service will be developed and adopted as a tool for applications to access EARTh and other specific domain thesauri.

6. SPECIFIC DOMAIN THESAURI

6.1. SnowTerm

SnowTerm is an example of a structured reference multilingual scientific and technical vocabulary, covering the terminology of a specific knowledge domain in the polar and the mountain environment. The thematic areas, covered at present, deal with snow and ice physics, snow and ice morphology, snow and ice radiometry, remote sensing and GIS applied to cryosphere environment, sea ice, avalanches, glaciers.

At present the database contains around 3.700 terms. Concerning the vertical structure the Classification Scheme already in use for the development of the EARTh Thesaurus was adopted [Plini et al., 2006].

6.2 GIS & RS Thesaurus

The project was started aiming at developing a controlled and structured terminology system related to GIS and remote sensing.

The first phase corresponded to the identification of potential sources of terminology, the selection of proper terms, their extraction and the creation of a terminology database.

Terms have been classified according to the EARTh upper structure. The semantic classification foresees the allocation of each term in a relation tree starting from the more general concept represented by the category.

The result is a English-Italian terminology system containing about 3.000 terms.

The system can be considered as a prototype of domain semantic map, open to future integrations and updating.

The use of the system by a user community would allow a better management of information and ease the information and data sharing, due both to the use of a common language and to a better semantic comprehension of data [Grignetti et al, 2005].

6.3. Pollution & Climate Changes Thesaurus

As a follow up of a project aiming at defining an action plan for the City of Rome that provides a concrete contribution to achieving the Kyoto Protocol targets in the set timeframe, we started the development of a terminology system on pollution and climate changes.

The resulting thesaurus contains at present around 1.500 terms in English and Italian. Being a very crucial topic it was decided to extend the collection of terminology in the direction of the interaction between pollution, pollutants and health. This activity would lead not only to an increase of the number of terms, but also to a consistent improvement of the associative relations. This work is currently ongoing.

ACKNOWLEDGEMENTS

We would like to express our gratitude to our former colleagues Fulvio Mazzocchi, Melissa Tiberi and Barbara De Santis, for their essential theoretical and practical contribution on the refinement of thesaurus relations and on multilingualism aspects.

REFERENCES

Barite, M. G., The Notion of ‘Category’. Its Implications in Subject Analysis and in the Construction and Evaluation of Indexing Languages, *Knowledge Organization*, Vol. 27 (2), 2000.

Bierwisch, M. and Kiefer F., Remarks on Definitions in Natural Language, in F. Kiefer (ed.), *Studies in Syntax and Semantics.* Dordrecht: Mouton, 1970.

Dahlberg, I., Conceptual Structures and Systematization, in *Categorie, oggetti e strutture della conoscenza.* Rome: CNR-ISRDS, 1995.

Fisher, D. H., From Thesauri towards Ontologies? *Advances in Knowledge Organization*, Vol. 6, 1998.

Fugmann, R., *Subject Analysis and Indexing.* Frankfurt: INDEKS Verlag, 1993.

Greenberg, J. et al., Final Report to the ALCTS/CCS Subject Analysis Committee, *Americam Library Association Report,* 1997. <www.ala.org/ala/alctscontent/catalogingsection/catcommitees/subjectanalysis/subjectrelations/finalreport.htm>.

Grignetti A., Plini P., Mazzocchi F., De Santis V., A Thesaurus for Remote Sensing and GIS: preliminary version and future plans. Proceedings of EnviroInfo 19th International Conference, Brno, 783-787, 2005.

Hudon, M., Multilingual thesaurus construction: Integrating the view of different cultures in one gateway to knowledge and concepts, *Knowledge Organization*, Vol. 24 (2), 1997.

Jackendoff , R., *Semantics and Cognition.* Cambridge, Mass: MITT Press, 1983.

Katz, J.J. and Fodor J.A., The structure of a semantic theory, *Language*, 39, 1963.

Kosovac, B., Internet/Intranet and Thesauri, *Project report - BELCAM "Roofing Thesaurus".* NRC Canada, Institute for Research in Construction, 1998. <http://irc.nrc-cnrc.gc.ca/thesaurus/roofing/report_b.html>.

IFLA Working Group on Guidelines for Multiligual Thesauri, *Guidelines for Multilingual Thesauri* - Draft, The Hague, International Federation of Library Associations and Institutions (IFLA), 2005, retrieved 20.01.2009 from the World Wide Web: http://www.ifla.org/VII/s29/pubs/Draft-multilingualthesauri.pdf

International Organization for Standardization (ISO), ISO 704: *Terminology work – Principles and methods,* 2nd ed. Geneva: ISO, 2000.

Mazzocchi F., De Santis B., Tiberi M., Plini, P., Relational Semantics in thesauri: Some Remarks at Theoretical and Practical Levels. *Knowledge Organization*, Vol. 34 (4), 197-214, 2007.

Milstead, J., Thesaurus Development: What Have We Learned? In *Multites Conference.* London: September 29-30, 2003. <www.multites.com/conference03.htm>.

Plini P., Salvatori R., Valt M., De Santis V., SnowTerm: a terminology database on snow and ice. Proceedings of the 21st Polar Libraries Colloquy. Rome, 82-89, 2006.

Schmitz-Esser, W., Thesaurus and Beyond: An Advanced Formula for Linguistic Engineering and Information Retrieval, *Knowledge Organization*, Vol. 26, No. 1, 1999.

Soergel, D. et al., Reengineering Thesauri for New Applications: the AGROVOC Example, *Journal of Digital Information*, Vol. 4, Issue 4, Article No. 257, 2004. http://jodi.ecs.soton.ac.uk/Articles/v04/i04/Soergel/.

Trigari, M., Old problems in a new environment. The impact of the Internet on multilingual thesauri as research interfaces, in *Multites Conference*. London: September 29-30, 2003. <www.multites.com/conference03.htm>.

Tudhope, D., Alani, H. and Jones C., Augmenting thesauri relationships: Possibilities for Retrieval, *Journal of Digital Information*, Vol. 1, Issue 8, Article No. 41, 2001. <http://jodi.ecs.soton.ac.uk/Articles/v01/i08/Tudhope/>.

Violi, P., *Significato ed esperienza*. Bompiani: Milano, 1997.

Wierzbicka, A., *Lexicography and Conceptual Analysis.* Ann Arbor: Karoma Publishers, 1985.

Winston, M. E., Chaffin, R. and Herrmann D., A Taxonomy of Part-Whole Relations, *Cognitive Science,* 11 , 1987.

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

Expert Cooperation in Forest Terminology: Services Offered by IUFRO

R. Prüller[a]

[a]*International Union of Forest Research Organizations (IUFRO), Hauptstrasse 7, 1140 Vienna, Austria (prueller@iufro.org)*

Abstract: Efficient communication among scientists and with decision makers requires a precise and clear technical language in forestry. The network of IUFRO experts in forest science is an immense pool of expert knowledge. Experts create the terms we use to communicate and thus constitute an ideal partner for IUFRO’s services in their attempt to define and make accessible the forest technical language to a wider public. At a time when English has become the *lingua franca* of science, it is crucial that other languages – IUFRO has four official languages – keep playing an active role. This diversity broadens the horizon of scientific thinking and knowledge, but also makes unequivocal communication more difficult to achieve. In order to ease communication in a multilingual world, we use vocabularies, glossaries and terminological databases which present the technical vocabulary and interpret it for the interested public. The role of forest experts is to assist in this process of explication of their own specialised language, in order to guarantee the liability of these communication tools. IUFRO offers the following services with regard to terminological problems through its Working Unit 6.03.02 “Trends in Forest Terminology” and its terminology project SilvaVoc based in the IUFRO Secretariat: Information platform; Electronic discussion groups; On-line Bibliography of terminological publications in forestry; Specialised glossaries, e.g. on carbon-related forest terminology, forest genetic resources, forest health; Multilingual terminological database SilvaTerm. Traditionally IUFRO’s role in terminology has been to make people aware of terminological differences. Our approach is therefore based more on descriptive than prescriptive principles, e.g. instead of aiming primarily at recommending definitions, it is our concern to point out differences in the use of the terms. IUFRO wishes to place its infrastructure of terminology work at the disposal of the international forestry community and invites the users to cooperation and calls for input and feedback.

***Keywords**:* Expert cooperation; Forest terminology; Terminological database

1. INTRODUCTION

The International Union of Forest Research Organizations (IUFRO) is the only world-wide international organization devoted to forest research and related sciences. It is a network of more than 700 members in over 110 countries which brings together research organizations, universities and individual scientists as well as decision-making authorities and other stakeholders with an interest and focus on forests and trees. IUFRO aims at exchanging and harmonizing research findings, communicating forest science within the network but also to the world. It understands itself as the advocate of forest science.

Terminology work has tradition in IUFRO. This paper describes the activities and services of the SilvaVoc project and extends the invitation to use and contribute to this clearinghouse on multilingual forest terminology.

2. SILVAVOC: STARTING POINT AND BASIC PROBLEMS

The need for continuing high quality work in forest terminology was expressed at the time of the 1990 IUFRO World Congress in Montreal. Subsequently IUFRO initiated a pilot project on forest terminology, SilvaVoc, in 1995 at the IUFRO Headquarters in Vienna.

The need of updating terminology data which existed within IUFRO was a starting point for terminology initiatives in IUFRO and at the same time the wish to raise awareness at international level of problems related to the technical language in forestry. Some of the following problems had to be addressed:

2.1 Inconsistency of concepts and definitions

Inconsistent use of forest concepts in conventions, international processes and organizations both on national and international level is a serious problem which causes confusion, for example when reporting on forest status or trends. The diversity of definitions is due to many factors, including the variety of ecological conditions, the differences of the practical management of forests and the way in which forest administration is organized by public institutions, and reflects the wide range of perceptions by various parts of society and their evolvement over time. Thus, there are dozens of definitions even for the most basic concepts, such as forest and tree.

2.2 Concept change

Moreover, concepts change over time. A change in society's conditions and perceptions often leads to broadening or narrowing of concepts. For example, it is quite obvious that the German concept of *Nachhaltigkeit* (sustainability) of the 19th century does not correspond with today's understanding of sustainability in forestry.

2.3 New terms (neologisms)

With the emergence of new scientific problems and research themes of changing priorities, new concepts and terms appear and come into use in forestry. These can be new creations like *eco-certification*, or general terms that have obtained a more technical interpretation, e.g., *mitigation* or *carbon sinks*. Especially terms of the latter type may seem obvious at first glance but for improved understanding and consistent use require discussion and definition, as well as clarification in a multilingual context.

2.4 Non-native knowledge of English

English is more and more used as the *lingua franca* of science. For a great number of scientists, however, English is not the mother tongue. In international meetings, participants are often confronted with a varying knowledge of English and specialized terminology.

3. IUFRO AND SILVAVOC TERMINOLOGY SERVICES

IUFRO's Special Project on Multilingual Forest Terminology SilvaVoc aims at fostering improved communication in forestry with special emphasis on multilingual aspects. Its objectives are :

- to provide bibliographical and consultation services about forest terminologies and terminological activities;
- to contribute to the harmonization and integration of existing and future terminological data;
- to make compiled and improved data available in electronic form through computer networks;
- to produce special glossaries for specific target groups in an appropriate medium of publication.

SilvaVoc relies on the scientists of the IUFRO structure which are gathered in eight Divisions (representing subject areas like silviculture, forest health, sustainable forest management etc.) and subdivided in 56 Research Groups and more than 150 Working Parties.

3.1 Cooperation with 6.03.02

A special cooperation links SilvaVoc and the IUFRO Working Party 6.03.02 Trends in Forest Terminology. In the past, several activities were designed in cooperation:

3.1.1 Electronic Discussion Groups

Discussion groups on specific issues, e.g. concepts needing clarification, or definitions needing standardization, are operated via Internet as moderated mailing lists. Their lifespan was determined by the moderator but should remain short, i.e. less than six months. The following concepts were explored in the past: old growth forests, natural forests, primary forests and related concepts; reforestation, afforestation and deforestation; low forest cover; and forest health. With the improved IUFRO homepage, this facility has now been transferred form the Working Party to the IUFRO HQ.

3.1.2 Terminological Awareness

In order to increase awareness for terminological issues among IUFRO members, IUFRO meeting organizers are called upon to clarify and discuss their terminology before or at the beginning of their meeting. The discussion should focus, e.g. on identifying recommended vs. non-recommended terms, on defining key terms of a specific field, or on pointing out differences in definitions between languages. The resulting information can be entered in the SilvaTerm database.

3.1.3 On-line Directory of Experts

This Directory of Experts had been on-line for a period of time, and provided the user with names of forestry experts who were willing to answer terminological questions in their field of expertise. Questions and answers were exchanged on a person to person basis. The users of the database were requested to send a short feedback to the database coordinator. Despite expression of continued interest this service was unfortunately not used as often as expected by the authors, and therefore had to be suspended.

3.2 SilvaVoc Information Platform

SilvaVoc provides a clearinghouse on the web with various services including entry points to:

3.2.1 International Bibliography of Forest Dictionaries

Documentation being the basis for subsequent work on terms and definitions, this on-line collection provides information on dictionaries, glossaries and terminological publications in many languages. All bibliographical entries are listed according to the Decimal Classification and specify the language(s) of the publication, the existence of equivalents, indexes, details on the arrangement of the terms, and if existing, links to an on-line version of the publication.

3.2.2 Specialised Glossaries

In cooperation with IUFRO Working Parties and FAO subject-specific glossaries have been established and are presented on-line according to different style types: -> Carbon glossary discusses the most recent terms and definitions related to carbon-related forest terminology; -> Forest genetic resources terminology has been designed together with forest genetists from FAO and IUFRO; and -> forest health glossary was elaborated in a project for an e-learning environment.

Additional features like entries to IUFRO In-house terminology, Information on other terminology projects, and Archives (comprising past activities, and papers and presentations) complete the SilvaVoc platform.

4. SILVAVOC ACTIVITIES IN MULTILINGUAL FOREST TERMINOLOGY

SilvaVoc has paid particular attention to the multilingual aspects of terminology work in forestry. It has acquired international recognition as a clearinghouse for multilingual forest terminology.

SilvaVoc participated in the international harmonization process of forestry concepts and terms as used by international conventions, processes and organisations. In order to foster common understanding of the definitions and the context in which they are developed, a series of Expert Meetings on Harmonizing Forest-Related Definitions for Use by Various Stakeholders was organized jointly by FAO, in collaboration with IUFRO, ITTO, the Intergovernmental Panel on Climate Change (IPCC), the Centre for International Forestry Research (CIFOR), and the United Nations Environment Programme (UNEP).

The First and Second Expert Meetings in 2002 concentrated on harmonizing core definitions of forest, afforestation, reforestation, and deforestation, and the Third Expert Meeting in January 2005 dealt with definitions related to the naturalness of forests, planted forests, trees outside forest, and forest management. It also focused on multilingual aspects and approved the continuation of the multi-stakeholder process under the Collaborative Partnership on Forests (CPF) umbrella. Harmonization – unlike standardization – can include adjustments for improved compatibility and consistency, establishing common elements, linkages and relationships between concepts in different languages, as well as documenting differences. The resulting comparative analytical framework for core forest-related terms (FAO 2002, FAO 2005) can be used to effectively compare definitions, thereby improving communication.

5. SILVATERM – IUFRO's TERMINOLOGY DATABASE

The multilingual SilvaTerm database is a central activity of IUFRO's terminology project. The database structure was designed in 1998 in cooperation with Professor Gerhard Budin, University of Vienna. The database is on-line since 1999, and it specializes exclusively in forestry and related areas. In 2006 the terminology database SilvaTerm was integrated into the IUFRO web management system.

At present, there are around 10,000 terminological records in SilvaTerm covering 50 forestry subject areas, from traditional fields like silviculture and management planning to wood structure, wee and pest control and meteorology to more recent research areas like urban forestry, remote sensing and recreation.

The following is a list of subject areas with an indicative number of entries per field:

anatomy	266	phenology	38
biology	128	physiography	34
botany	102	physiology	152
cable logging	173	pulp and paper	299
chemistry	263	range management	176
diseases	172	recreation	91
ecology	827	remote sensing	157
economics	123	roads	74
engineering	352	sawing	361
entomology	199	seasoning	260

fire management	284	silviculture	823
forest management	350	soil sciences	505
genetics	495	sorting	177
gluing	296	statistics	306
harvesting	244	survey	148
hydrology	205	transport	131
land management	131	weed and pest control	165
machinery	340	wildlife	145
manufacturing	227	wood finishing	57
mechanics	295	wood preservation	181
mensuration	371	wood structure	477
meteorology	92	wood-based panels	481
mycology	107	work safety	154
other forest products	137		

Figure 1. Distribution of subject areas in the SilvaTerm database

A total of more than 9,000 terms is related to subject areas, an additional 400 terms with no subject area allocated include themes like carbon sequestration and climate change.

Almost all entries are multilingual, in English, French and Spanish, approximately 20% of all entries include additional equivalent terms in German, Italian, Portuguese, Japanese, Hungarian and Swahili. A special solution has been created for Japanese with a parallel database that is linked to SilvaTerm and provides Kanji character entries. More work is ongoing to include terms and definitions in Polish, Romanian, Czech and Russian.

Definitions are given for 40% of the English entries, and for 20% of the entries in French, Spanish and German. In general, internationally agreed definitions, in particular those developed by the FAO/CIFOR/IPCC/IUFRO/ITTO/UNEP Process on Harmonizing Forest-Related Definitions for Various Stakeholders, by International Conventions or other standardizing bodies have been given priority, and are presented with their reference source. In order to take account of the diversity of approaches, in some cases more than one definition is indicated with the respective source. Explanations of the differing use of terms are given in the explanatory notes.

The database records provide terms and equivalents, short forms, synonyms, definitions, use and technical notes, cross references, sources, the relevant IUFRO Units, subject areas and links to additional information on the Internet. These hyperlinks to additional information can be displayed for 10% of the English terms. A special feedback field calls for interaction, critical comments and information about missing or new entries.

Key data from glossaries produced by SilvaVoc on carbon in forests, genetic resources and the pocket glossary are presently incorporated.

6. CONCLUSION

Forests and forestry issues are now being discussed on a much broader basis and by more international policy fora than ever before. It is therefore of extreme importance to develop a common understanding of the concepts used in international discussions, e.g. the term degradation in the ongoing negotiations on Reducing Emissions from Deforestation and forest Degradation (REDD).

With its four official languages and long-lasting commitment to terminology, IUFRO is currently addressing how to best reconcile consensus and diversity. Despite funding problems for the terminology project SilvaVoc, IUFRO is still involved in harmonization processes with other international bodies, seeking common ground for critical terms such as *afforestation*, *reforestation* and *deforestation* or most recently *forest degradation*. With its global scope and network of forest scientists, it is perhaps the best resource for addressing international multilingual forest terminology issues while contributing to and benefiting from this unique pool of forest expertise.

In order to be able to provide multilingual services in forest terminology in the future, IUFRO-SilvaVoc counts on cooperation with forest experts within and outside IUFRO. Users of the services are considered to be contributors as well, and are therefore invited to provide information and thereby share terminological data with colleagues and forest scientists.

REFERENCES

FAO, Second expert meeting on harmonizing forest-related definitions for use by various stakeholders. Proceedings. Food and Agriculture Organization of the United Nations (FAO). Rome, 11–13 September 2002. http://www.fao.org/DOCREP/005/Y4171E/Y4171E00.HTM

FAO, Third expert meeting on harmonizing forest-related definitions for use by various stakeholders. Proceedings, Food and Agriculture Organization of the United Nations (FAO). Rome, 17-19 January 2005. ftp://ftp.fao.org/docrep/fao/008/j4959e/j4959e.pdf

SilvaVoc homepage: www.iufro.org/science/special/silvavoc

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pilmmann, I.Holoubek, R. Legat (eds.)
Masaryk University, 2009

EcoLexicon: A Frame-Based Knowledge Base for the Environment

Arianne Reimerink[a] and Pamela Faber[b]

[a] *Department of Translation and Interpreting, University of Granada, C/ Buensuceso 11, 18002, Granada, Spain (arianne@ugr.es)*

[b] *Department of Translation and Interpreting, University of Granada, C/ Buensuceso 11, 18002, Granada, Spain (pfaber@ugr.es)*

Abstract: EcoLexicon (http://manila.ugr.es/visual) is a frame-based specialized knowledge resource on the environment. Although currently it can be regarded as a linguistic ontology, it is gradually evolving towards the status of a formal ontology. As such, it focuses on: (1) conceptual organization; (2) the multidimensional nature of terminological units; (3) the extraction of semantic and syntactic information through the use of multilingual corpora. Based on Frame Semantics [Fillmore 1977, 1982, 1992], EcoLexicon is structured in terms of knowledge frames that are organized in an ontological structure. In this case, one of the objectives of EcoLexicon is to achieve interoperability for descriptions in Spanish, English, and German of environmental entities (concepts and roles). At a macrostructural level, the most generic conceptual categories are configured in a prototypical domain event which provides a frame for the organization of more specific concepts in sub-hierarchies [Faber et al. 2007], [León, Reimerink, Faber 2008]. At the microstructural level, specialized concepts are described in terms of a closed inventory of conceptual relations. EcoLexicon represents this conceptual structure of the environmental domain in a visual thesaurus. EcoLexicon and its multimodal visualization of the environmental knowledge domain combines conceptual information with linguistic data and images to meet the communicative as well as cognitive needs of different types of end-users such as, students, researchers, translators, technical writers and experts.

Keywords: Environmental domain; Specialized knowledge representation; Terminography; Frame; Ontology.

1. INTRODUCTION

EcoLexicon is a frame-based multilingual knowledge resource on the environment. It is a knowledge base because the information is structured and interrelated so as to form a network. This type of network can be regarded as an incipient linguistic ontology, whose objectives are to produce new knowledge, prove the consistency of existing knowledge, and to enhance searches. As such, it focuses on: (1) conceptual organization; (2) the multidimensional nature of terminological units; (3) the extraction of semantic and syntactic information through the use of multilingual corpora.

Figure 1 shows a screen capture of the EcoLexicon representation of the general concept of INSTRUMENT. For example, the top level categories in INSTRUMENT represent the principal types of function that a scientific instrument can have in this domain through its subtypes: (i) measuring instrument; (ii) recording instrument; (iii) sampling instrument. The relations here include the fact that all these concepts are a type of instrument and the different instances of each class. Clicking on any other concept on the screen accesses the visualization of a new network, centered on the new concept and its conceptual relations.

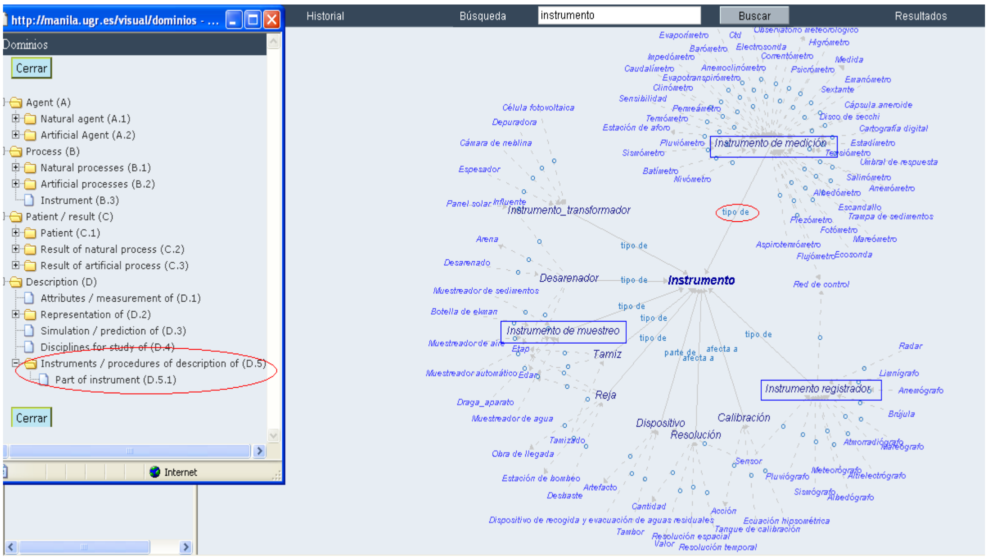


Figure 1. EcoLexicon ThinkMap representation: INSTRUMENTO [INSTRUMENT]

The theory underlying this resource is called *frame-based terminology* [Faber et al. 2005, Faber et al. 2007, Faber et al. 2008]. Frame-based terminology uses a modified version of Fillmore's frames [Fillmore 1977, 1982, 1985, Fillmore and Atkins 1992] coupled with premises from Cognitive Linguistics to structure specialized domains and create non-language-specific representations. The main focus is on conceptual relations as well as a term's combinatorial potential.

In frame-based terminology, conceptual networks are based on an underlying domain event [Faber and Jiménez 2002, Faber et al. 2007]. This prototypical domain event or action-environment interface [Barsalou 2003] provides a template applicable to all levels of information structuring. In this way a structure is obtained which facilitates and enhances knowledge acquisition since the information in term entries is internally as well as externally coherent [Faber et al, 2007].

The methodology used to derive the conceptual system of the domain integrates both a top-down and bottom-up approach. The bottom-up approach extracts information from a corpus of texts in various languages, specifically related to the domain. The top-down approach includes the information provided by specialized dictionaries and other reference material, complemented by the help of experts in the field.

2. MACROSTRUCTURE: ENVIRONMENTAL EVENT

In EcoLexicon, knowledge can be accessed from top-level categories to more fine-grained relational structures. The most generic level is the Environmental Event (EE), which provides a frame for the organization of more specific concepts. As shown in Figure 2 the EE is conceptualized as a dynamic process that is initiated by an agent (either natural or human). This process, affects a specific kind of patient (an environmental entity), and produces a result. These macro-categories (AGENT, PROCESS, PATIENT, etc.) are the concept roles characteristic of this specialized domain. Additionally, there are peripheral categories which include INSTRUMENTS that are typically used during the EE, as well as a category where the concepts of measurement, analysis, and description of the processes in the main event are included.

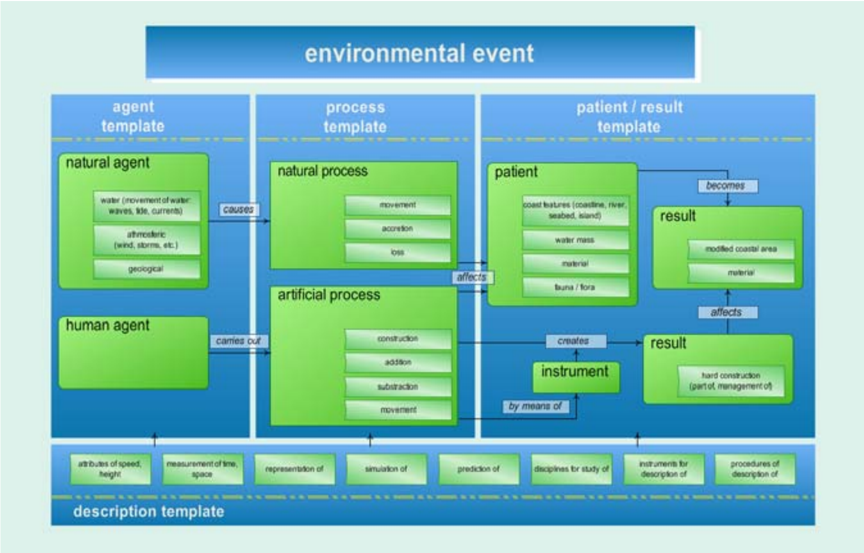


Figure 2: Environmental Event (http://manila.ugr.es/visual/info.html)

3. MICROSTRUCTURE

3.1 Conceptual relations

The primary conceptual relations that link semantic roles and the macro-categories (AGENT, PROCESS and PATIENT/RESULT) are non-hierarchical relations such as AFFECTS, CAUSES, and RESULTS. To represent the dynamicity of the environmental domain, more specific concepts within each macro-category are organized in a hierarchical network where they are linked by both hierarchical and non-hierarchical relations. In our research in the field we have found the following inventory of relation types, some of which are domain-specific: ISA, PART_OF, MADE_OF, DELIMITED_BY, LOCATED_AT, TAKES_PLACE_IN, ATTRIBUTE_OF, RESULT_OF, AFFECTS, EFFECTED_BY, HAS_FUNCTION, MEASURES, REPRESENTS and STUDIES.

3.2 Concept definitions

The definition of each concept is systematically related to its place in the conceptual system. This makes all levels complementary as well as highly consistent. Our definitions accomplish the following: (1) they make category membership explicit; (2) they reflect a concept's relations with other concepts within the same category; (3) they specify essential attributes and features.

As a result, definitions are based on definitional templates shared by all concepts within the same conceptual category. For example, the definition of all PROCESSES includes information regarding their superordinate category.

PROCESS: sequence of actions or steps that happen or take place.	**ARTIFICIAL PROCESS:** **process** taken in order to achieve a particular end (with an instrument).
• ISA • (PART_OF)	• ISA • (PART_OF) • HAS_FUNCTION • (EFFECTED_BY)

Figure 3. PROCESS and ARTIFICIAL PROCESS templates [León, Reimerink, Faber 2008]

Figure 3 shows the definitional template of the concept of PROCESS, which is composed of the ISA and PART_OF relations. As shown, ARTIFICIAL PROCESS contains this same information as well as two other relations (HAS-FUNCTION and EFFECTED-BY). These two relations reflect the fact that artificial processes are initiated by a human agent for a reason and generally entail an instrument.

The definition of the artificial process of DEPURATION, is specifically based on this template. The resulting definition is consistent with other artificial process definitions because it marks category membership and specifies dimensions common to all the concepts within the category. Depuration has different phases, and can be carried out with a wide variety of methods and instruments. Since the process affects waste water, the dimension HAS-PATIENT is added.

PURIFICATION: artificial process that eliminates contaminants from waste water by mechanical, biological, or physical-chemical methods.
• Artificial process [ISA] • Sedimentation, etc. [PART_OF] • Eliminate [HAS_FUNCTION] • Mechanical methods [EFFECTED_BY] • Waste water [HAS_PATIENT]

Figure 4. Activation of the definitional template of ARTIFICIAL PROCESS in PURIFICATION

4. ECOLEXICON VISUAL THESAURUS

4.1 Conceptual information in EcoLexicon entries

As previously mentioned, EcoLexicon is a knowledge resource that integrates a wide variety of information regarding environmental concepts. As can be observed in Figure 5, PURIFICATION appears conceptually framed at three different levels. It has the macro-role of subtraction within the EE because it eliminates contamination (Subtraction B.2.3). Secondly, it appears related to other conceptual networks through different relation types, such as ISA, AFFECTS and EFFECTED_BY (marked in red). Finally, a linguistic description, based on the template of the previous section, of the concept is given (the white square).

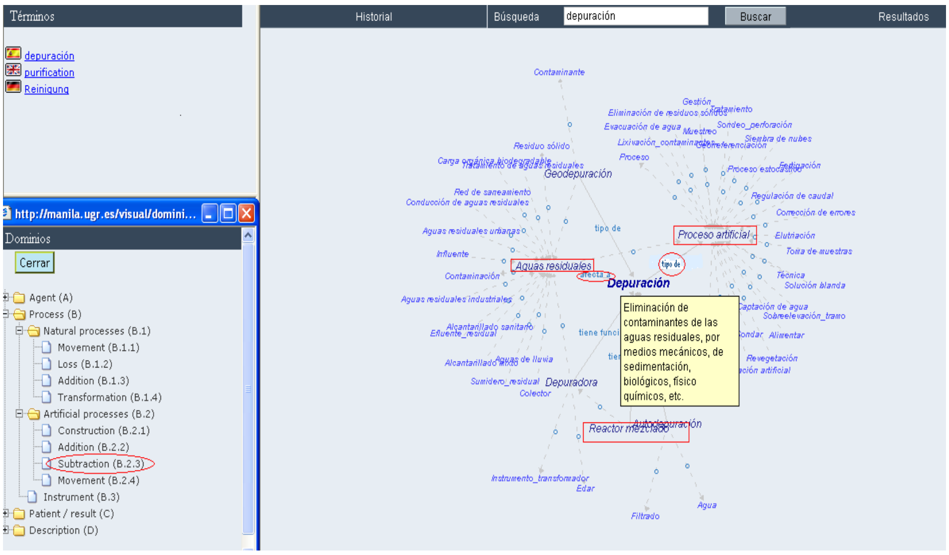


Figure 5. Thinkmap representation: DEPURACIÓN [PURIFICATION]

The interface currently uses Spanish for the language-independent conceptual information. To reach a wider public, an English version of this language-independent structure will be provided shortly.

4.2 Terms: Contextual information

Apart from conceptual information in the definition and the conceptual relational organization, each entry includes the English, Spanish, and German terms linked to the concept. The description of each term includes basic grammatical information such as gender, number, etc. as well as contextual information. This information shows how the term functions in a text. We have included three types of contextual information: contexts, concordances and images.

Contexts

A knowledge-rich context (Meier 2001) in a knowledge base is one that helps the user achieve a level of understanding of (part of) a specialized domain. A context should always add something to the definition to avoid redundancy. New information can be added by: (1) giving more detailed information on one of the relations expressed in the definition; (2) providing information on other relations than the ones expressed in the definition.

An example of how contexts can be selected can be found in the term entry for the concept DREDGER. A dredger is a machine used for dredging operations and as such must be directly linked to the concept DREDGING in the terminological resource. The definitional template of this object and the related process are the following:

DREDGER
Instrument [ISA] used for dredging operations [HAS_FUNCTION].

DREDGING/DREDGE
Removal of beach material from underwater [ACTION] by pumping, extracting and piping it [PHASES] by means of a dredger [EFFECTED_BY] in order to maintain water depths in rivers, canals or harbours and to obtain material for construction or beach nourishment [HAS_FUNCTION].

Figure 6. Definitions and relations of DREDGER and DREDGE

Example (1) is a context that contains several elements of domain knowledge, it gives a definition of the process DREDGING and it explains that this process is carried out by means of a specific instrument.

(1) Dredging is the operation of removing material from under water. In all but a few situations the excavation is undertaken by a specialist floating plant, known as a dredger.

The content directly relates the instrument with the function that it carries out, and therefore, seems to be an appropriate context for a terminological resource. However, the relations expressed in the context [ISA] and [HAS_FUNCTION] are the basic relations expressed in the definition of both concepts DREDGER and DREDGING. Adding this context to the entry in the database does not add any new information.

Therefore, the context in (2) is more appropriate for database construction since it activates the basic relations between both concepts.

(2) Dredging can be carried out wherever there is sufficient water depth to allow a dredger to operate. Thus dredgers can be found around the coasts, in rivers and in canals. They can be found in lakes and ponds far from the sea and they can be found in exposed offshore locations far from land.

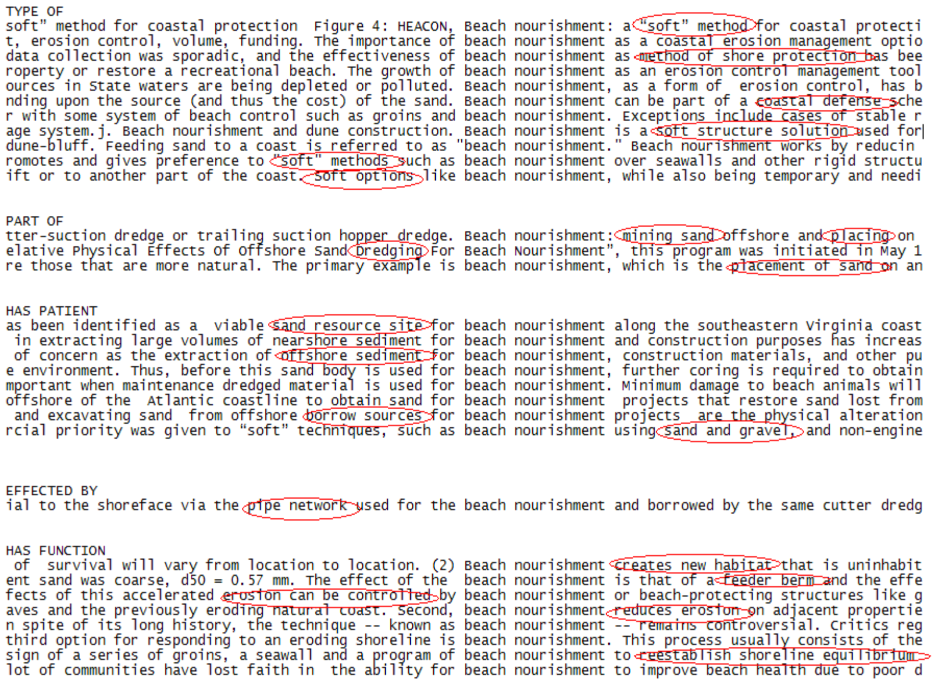
The context in (2) focuses on the possible location of the instrument. Location is not one of the core knowledge elements of the definitional template of dredger, but it is present in the definitional template of dredge: *from underwater*. The context, therefore, gives the user additional information.

Concordances

BEACH NOURISHMENT: soft coastal defence action consisting of replenishing a beach with new dredged materials in order to protect the coastline from erosion, storms and floods.

- Soft coastal defence action [ISA]
- Replenishing, dredging [PART_OF]
- Protect from erosion, storms and floods [HAS_FUNCTION]
- Dredger [EFFECTED_BY]
- Coastline [HAS_PATIENT]
- Dredged materials [HAS_PATIENT]

Figure 7. Linguistic activation of the category structure of BEACH NOURISHMENT



```
TYPE OF
soft" method for coastal protection  Figure 4: HEACON, Beach nourishment: a "soft" method for coastal protecti
t, erosion control, volume, funding. The importance of beach nourishment as a coastal erosion management optio
data collection was sporadic, and the effectiveness of beach nourishment as method of shore protection has bee
roperty or restore a recreational beach. The growth of beach nourishment as an erosion control management tool
ources in State waters are being depleted or polluted. Beach nourishment, as a form of  erosion control, has b
nding upon the source (and thus the cost) of the sand. Beach nourishment can be part of a coastal defense sche
r with some system of beach control such as groins and beach nourishment. Exceptions include cases of stable r
age system.j. Beach nourishment and dune construction. Beach nourishment is a soft structure solution used for|
dune-bluff. Feeding sand to a coast is referred to as "beach nourishment." Beach nourishment works by reducin
romotes and gives preference to "soft" methods such as beach nourishment over seawalls and other rigid structu
ift or to another part of the coast. soft options like beach nourishment, while also being temporary and needi

PART OF
tter-suction dredge or trailing suction hopper dredge. Beach nourishment: mining sand offshore and placing on
elative Physical Effects of Offshore Sand Dredging For Beach Nourishment", this program was initiated in May 1
re those that are more natural. The primary example is beach nourishment, which is the placement of sand on an

HAS PATIENT
as been identified as a  viable sand resource site for beach nourishment along the southeastern Virginia coast
 in extracting large volumes of nearshore sediment for beach nourishment and construction purposes has increas
 of concern as the extraction of offshore sediment for beach nourishment, construction materials, and other pu
e environment. Thus, before this sand body is used for beach nourishment, further coring is required to obtain
mportant when maintenance dredged material is used for beach nourishment. Minimum damage to beach animals will
offshore of the  Atlantic coastline to obtain sand for beach nourishment  projects that restore sand lost from
 and excavating sand  from offshore borrow sources for beach nourishment projects  are the physical alteration
rcial priority was given to "soft" techniques, such as beach nourishment using sand and gravel, and non-engine

EFFECTED BY
ial to the shoreface via the pipe network used for the beach nourishment and borrowed by the same cutter dredg

HAS FUNCTION
 of  survival will vary from location to location. (2) Beach nourishment creates new habitat that is uninhabit
ent sand was coarse, d50 = 0.57 mm. The effect of the  beach nourishment is that of a feeder berm and the effe
fects of this accelerated erosion can be controlled by beach nourishment or beach-protecting structures like g
aves and the previously eroding natural coast. Second, beach nourishment reduces erosion on adjacent propertie
n spite of its long history, the technique -- known as beach nourishment -- remains controversial. Critics reg
third option for responding to an eroding shoreline is beach nourishment. This process usually consists of the
sign of a series of groins, a seawall and a program of beach nourishment to reestablish shoreline equilibrium
lot of communities have lost faith in  the ability for beach nourishment to improve beach health due to poor d
```

Figure 8. Conceptual concordances: BEACH NOURISHMENT [León, Reimerink, Faber 2008]

Our knowledge base establishes three types of concordances: conceptual, phraseological and verbal. In this way, users can widen their linguistic knowledge of a term from different perspectives. For example in Figure 8, the concordances of the process BEACH NOURISHMENT show the same dimensions represented in its terminographic definition (TYPE_OF, PART_OF, HAS_PATIENT, EFFECTED_BY and HAS_FUNCTION) in Figure 7. It appears as a type of *soft method, coastal defence option, soft solution,* or *method of shore protection*. These are different designations to refer to the same superordinate concept SOFT COASTAL DEFENCE ACTION, which also provide users with a wider knowledge on linguistic variation in specialized discourse.

Graphical information

Depending on the concept, graphical information can often be useful to enhance knowledge representation. Linguistic and graphical descriptions of specialized entities play a major role in the understanding process when both types of description converge to highlight the multidimensional nature of concepts as well as the set of conceptual relations typical of a specific subdomain. When the information in such descriptions is enhanced with images that highlight the different types of conceptual information within each definition, the entry is substantially enriched.

In EcoLexicon images are classified, according to their most salient functions [Anglin et al. 2004, Faber et al. 2007] in terms of their relationship with the real-world entity represented. Image classification is thus based on the criteria of iconicity, abstraction and dynamism:

- Iconic images are especially useful for the description of physical objects, as they resemble the real-world object represented through the abstraction of conceptual attributes in the illustration.
- Abstraction in an illustration is a matter of degree, and refers to the cognitive effort required for the recognition and representation of the concept [Levie and Lentz 1982, Park and Hopkins 1993, Rieber 1994]. A higher level of abstraction is therefore more applicable to the description of mental objects.
- Dynamism implies the representation of movement. Video and animation, as well as images showing different phases or steps, are the best choice when describing processes.

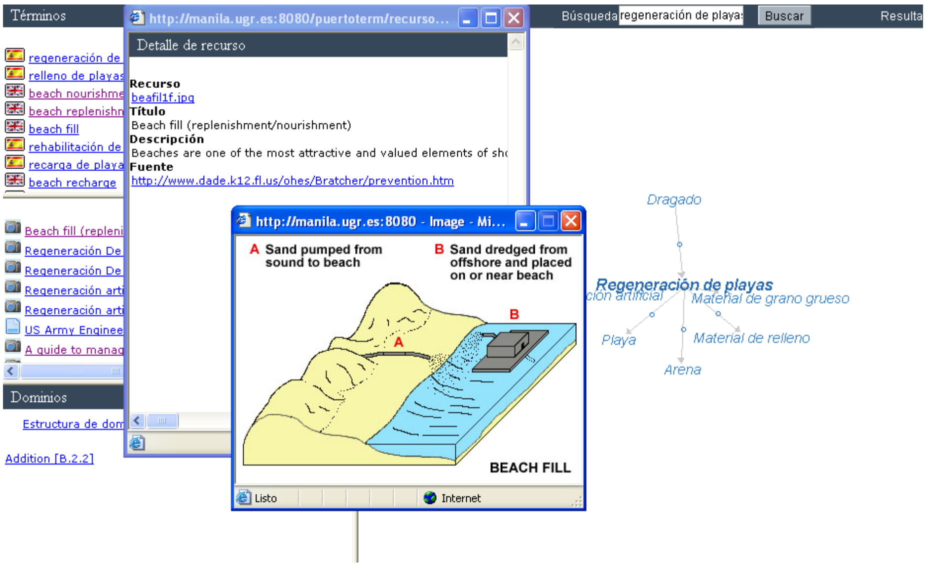


Figure 9. Image representing REGENERACIÓN DE PLAYAS [BEACH NOURISHMENT]

The concept BEACH NOURISHMENT is a process; therefore it can be understood as a sequence of phases which are part of that process. In this case, the ISA and PART_OF relations are encoded by images with a certain degree of abstraction. In the knowledge base, all images are selected according to the relations expressed by the definitional template of the concept. The user can visualize these images (see Figure 9) by clicking on the link in the resources (*recursos*; in the left margin) related to the concept.

5. CONCLUSIONS

This paper describes the theoretical and methodological bases of EcoLexicon, a conceptually-structured knowledge base, providing conceptual, linguistic, and graphical information regarding environmental concepts. One of the primary considerations in creating this knowledge base was that an adequate knowledge resource should not be a mere repository for bits and pieces of information that have been scavenged from different sources, and gathered together in a kind of conceptual magpie’s nest. This meant that all elements in the data fields of concept entries had to be justified and coherently liked to each other.

The organization of EcoLexicon is thus geared to providing the user with an understanding of the meaning of a concept as well as sufficient knowledge of its location within the knowledge structure of the domain. For this reason, conceptual representations contain information in various formats. Linguistic and graphical descriptions of specialized entities also play a major role in knowledge representation, especially when both converge to highlight the multidimensional nature of concepts as well as the conceptual relations within a specialized domain. As shown in this article, thanks to conceptual templates characteristic of each category, the information in EcoLexicon concept definitions reflect network structure, and mesh with visual information in images for a better understanding of complex and dynamic concept systems.

ACKNOWLEDGEMENTS

This research is part of the project ECOSISTEMA: *Espacio úniCO de SIStemas de información ontológica y TEsaurus sobre el Medio Ambiente* (FFI2008-06080-C03-01/FILO), funded by the Spanish Ministry of Science and Innovation

REFERENCES

Anglin, G. Vaez, H. and K, Cunningham, Visual Representations and Learning: the Role of Static and Animated Graphics, in Jonassen, D. *Handbook of Research on Educational Communications and Technology*. New Jersey: Lawrence Erlbaum Associates, 2004.

Barsalou, L.W., Situated Simulation in the Human Conceptual System, *Language and Cognitive Processes* 18, 513-562, 2003.

Faber, P and C. Jiménez Hurtado (eds.), *Investigar en terminología*. Granada: Comares, 2002.

Faber, P., C. Márquez Linares and M. Vega Expósito, Framing terminology: a process-oriented approach, *Meta* 50(4), CD-ROM, 2005.

Faber, P., P. León Araúz and J. A. Prieto Velasco and A. Reimerink, Linking images and words: the description of specialized concepts, *International Journal of Lexicography* 20, 39-65, 2007.

Faber, P., P. León, J. A. Prieto, Semantic relations, dynamicity, and terminological knowledge bases, *Proceedings of the XVIII FIT World Congress*, CD-ROM, Shanghai: FIT, 2008.

Fillmore, C. J., Scenes-and-frames semantics, in Zampolli, A. (ed.), *Linguistics Structures Processing*, Amsterdam/New York: North Holland Publishing Company, 55-81, 1977.

Fillmore, C. J., Frame semantics, in the Linguistic Society of Korea (ed.), *Linguistics in the Morning Calm*, Soeul: Hanshin, 111-137, 1982

Fillmore, C. J., Frames and the semantics of understanding, *Quaderni di Semántica*. 6 (2), 222-254, 1985.

Fillmore, C. J., Starting where the dictionaries stop: the challenge of corpus lexicography, in Atkins, B. T. S. and A. Zampolli (eds.), *Computational Approaches to the Lexicon*, Oxford: Oxford University Press, 349-393, 1994.

Fillmore, C. J. and B. T. Atkins, Towards a frame-based organization of the lexicon: the semantics of RISK and its neighbours, in Lehrer, A. and E. Kittay (eds.), *Frames, Fields, and Contrasts: New Essays in Semantics and Lexical Organization*. Hillsdale: Lawrence Erlbaum, 75-102, 1992.

León Araúz, P., A. Reimerink, and P. Faber, PuertoTerm & MarcoCosta: a Frame-Based Knowledge Base of Coastal Engineering, *Proceedings of the XVIII FIT World Congress*, CD-ROM, Shanghai: FIT, 2008.

Levie, W. H. and R. Lentz, Effects of text illustrations: a review of research, *Educational Communication and Technology Journal*, 30, 195-232, 1982.

Meyer, I., Extracting knowledge-rich contexts for terminography: A conceptual and methodogical framework, in Bourigault, D., C. Jacquemin and M. C. L'Homme (eds), *Recent Advances in Computational Terminology*, Amsterdam: John Benjamins, 279-302, 2001.

Park, O. and R. Hopkins, Instructional Conditions for Using Animated Visual Displays: A Review, *Instructional Science,* 22, 1–24, 1993.

Rieber, L.P., *Computers, Graphics, and Learning*. Madison, WI: Brown & Benchmark, 1994.

Workshop 5

SEIS/NESIS - ICT Components for SEIS around Europe

Organized by Jarmila Cikánková**, Giorgio Saio,**
Danny Vandenbroucke and Rafal Wawer

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

PortalU®, a Tool to Support the Implementation of the Shared Environmental Information System (SEIS) in Germany

F. Kruse[a], **S. Uhrich**[a], **M. Klenke**[a], **H. Lehmann**[a], **C. Giffei**[a], **S. Töpker**[a]

[a] *Coordination Centre PortalU at the Lower Saxony Ministry of Environment and Climate Protection, Archivstr. 2, 30169 Hannover, Germany (kst@portalu.de)*

Abstract: The German Environmental Information Portal, PortalU®, provides one-stop access to government-owned environmental information in Germany. PortalU® gives access to more than 2,500,000 webpages and more then 500,000 data and metadata sets from different national, regional and local authorities. Organisationally PortalU® is the result of a cooperation of the federal government and the 16 German states. PortalU® is build up by the software InGrid®. The software is realised very modular. It includes several interfaces including WMS, CSW and OpenSearch. The portal has a set of viewing components for the visualisation of search results, maps and metadata content. A metadata catalogue, the InGrid®Catalog, is integrated in the software. Core component of this catalogue is the ISO- and INSPIRE-compliant InGrid®Editor for collection and maintenance of metadata. A concept for the exchange and visualisation of monitoring data on the basis of the OGC-defined sensor observation service (SOS) is under development. This concept could be a model for the data exchange in a European Shared Environmental Information System (SEIS). The integration of SOS-services for data exchange and visualisation is the next step of advancement of InGrid® and PortalU®. From an organisational as well as from a technical point of view PortalU® could be a prototype for a pan-European shared environmental information system.

Keywords: Environmental information; Environmental data; Environmental information space; Environmental information system.

1. INTRODUCTION

The German Environmental Information Portal, PortalU®, provides one-stop access to government-owned environmental information in Germany. The portal integrates the access to a large number of heterogeneous and geographically distributed information sources organisationally and technically. A user-friendly interface in addition to advanced search and visualization tools enable both experts and non-experts to find and view texts of national, regional and local legislation, information about environmental policies and programmes, environmental reports, monitoring data, digital maps, and many other types of environmental information and data. As an online information portal, PortalU® (www.portalu.de) is freely accessible to all internet users.

The portal features a number of information services, among them up-to-date environmental news, access to environmental monitoring data, chronicles of environmentally relevant events, and links to new publications and events (Figure 1). Most importantly, however, PortalU® maintains an index of environmental information (metadata catalog) held by public authorities in Germany. Today PortalU® gives access to more than

2,500,000 webpages and more then 500,000 data and metadata sets from different national, regional and local authorities. The portal functions as a highly-visible and central access point to this information.

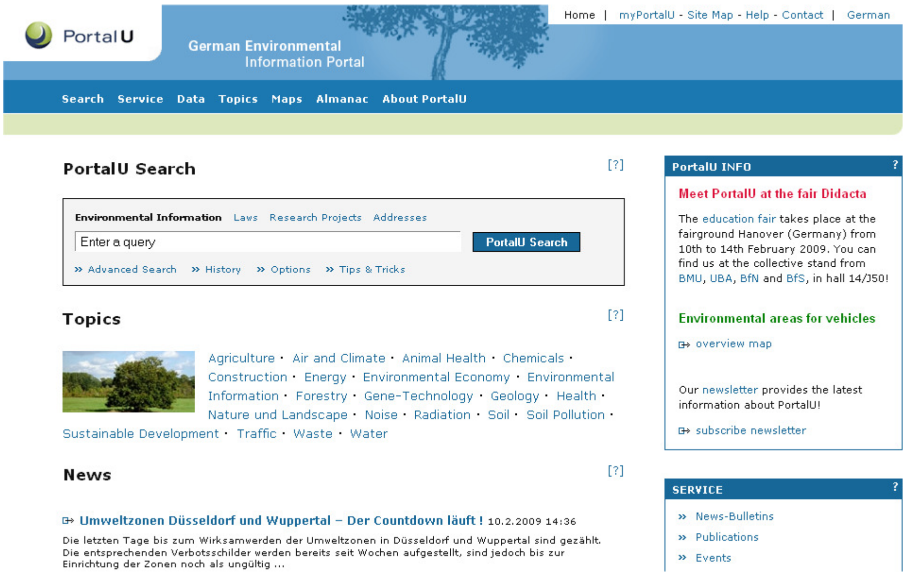


Figure 1: The German Environmental Information Portal – PortalU®

2. ORGANISATIONAL ASPECTS

Organisationally PortalU® is the result of a cooperation of federal and regional environment agencies. Funded through an administrative agreement between the federal government and the 16 German states (Länder), PortalU® could be established as a sustainable long-term project. This is important, because Germany is a federal structured country. Many information and data are collected and maintained by the regional level, the German states. So PortalU® is a national shared environmental information system and could be a model for a pan-European shared environmental information system (SEIS) not only in technical but also in organisational sense.

PortalU® is part of the administration's strategy to comply with the Aarhus-Convention and EU-Directive 2003/4/EC, both calling for better public access to environmental information. Furthermore it is the national information knot for environmental metadata in the context of the INSPIRE Directive. In future it will be the national implementation of the pan-European shared environmental information system (SEIS).

3. THE TECHNICAL STRUCTURE

The modular software of PortalU® is called InGrid®. It bundles the decentralized distributed environmental information. In InGrid® all environmental information web pages, databases as well as data catalogues can be searched by a powerful search engine. If required, spatially, temporally or semantically restrictions can be taken into account. The query results are presented in a main and a secondary result list. The results in the main list are ranked according to the relevance of the query terms for the indexed documents. The secondary list contains results from data bases, which are connected by an open interface. These results are sorted by provider. Besides the full text search, separately prepared environmental topics from A like “air and climate” over N like “nature and landscape” to W like “water” can be browsed. These thematic web pages are particularly relevant environmental information, selected by experts to improve the understanding of environmental information for non-experts. The 21 environmental topics are subdivided in six functional classes: legislation, concepts, reports, state-of-the-environment, data and maps as well as risk-assessment. Thus in the class “water – legislation” information about waste water legislations or the Water Framework Directive (2000/60/E) should be found. The environmental and functional classes comprehend the topics, which were mentioned in annex B of the EEID (2003/4/EG). Furthermore additional components are available to improve the understanding of environmental information: the rubric “service” provides press releases as well as information about publications and events of the connected public authorities and the rubric “data” provides environmental monitoring data.

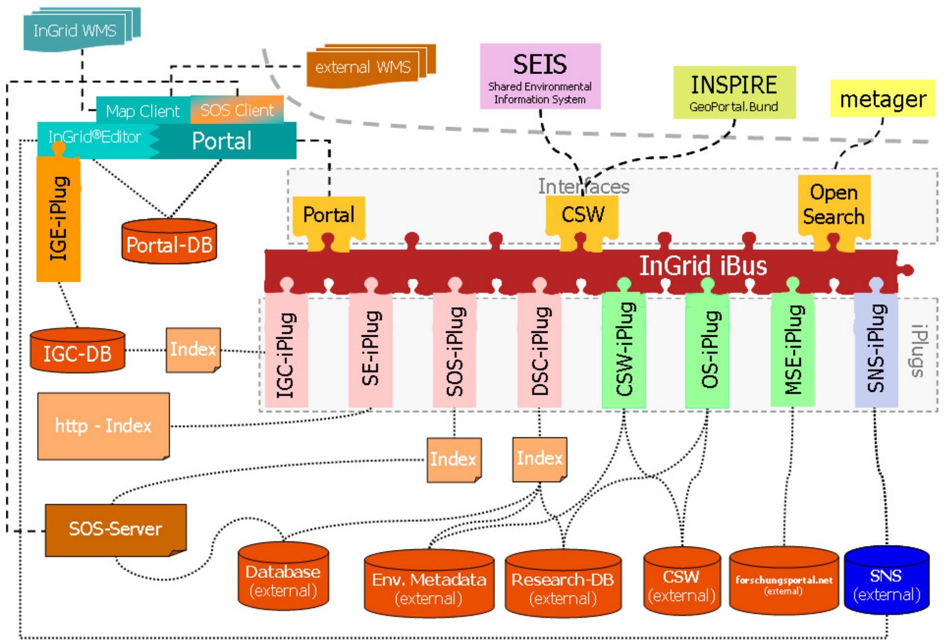


Figure 2: InGrid® Architecture

The PortalU® software InGrid® is structured in several components (Figure 2). The information bus (iBus) forms the central component of InGrid®. It receives and processes search queries, which come from the portal surface or from other connected interfaces. The search queries are transferred to the data sources and the query results were bundled and delivered again by the iBus. The information plugs (iPlugs) build a further part of InGrid®. They can be described as generic adapters to connect data sources to the iBus. Different iPlugs are implemented to connect different kinds of data sources. Data bases and expert information systems for instance are connected by the data source client (DSC) iPlug. Thus, an access to parts of the so called hidden web is realised. New data sources are easily integrable by connecting them to existing iPlugs or by adding a new specific iPlugs. The InGrid® database module consists of two parts: the database of the InGrid®Catalogue (IGC) and the portal database. Metadata of the environmental catalogues are stored in the IGC, while internal information like the user administration are stored and managed in the portal DB. Furthermore metadata can be created and managed with the InGrid®Editor (IGE). Further components are an integrated web map service (WMS), a WMS viewer and certainly the websurface of the portal.

4. THE INGRID®CATALOG

One of the core components of PortalU® is a metadata catalogue (InGrid®Catalog) for environmental information. The metadata model of the InGrid®Catalog is compatible to geographical standards like ISO 19115 and 19119 and the INSPIRE implementation rules and beyond these it is quite flexible. It allows storing of meta-information not only for geographic datasets and series and for geographic services but also for all kinds of environmental information such as non-geographic datasets, non-geographic services and documents.

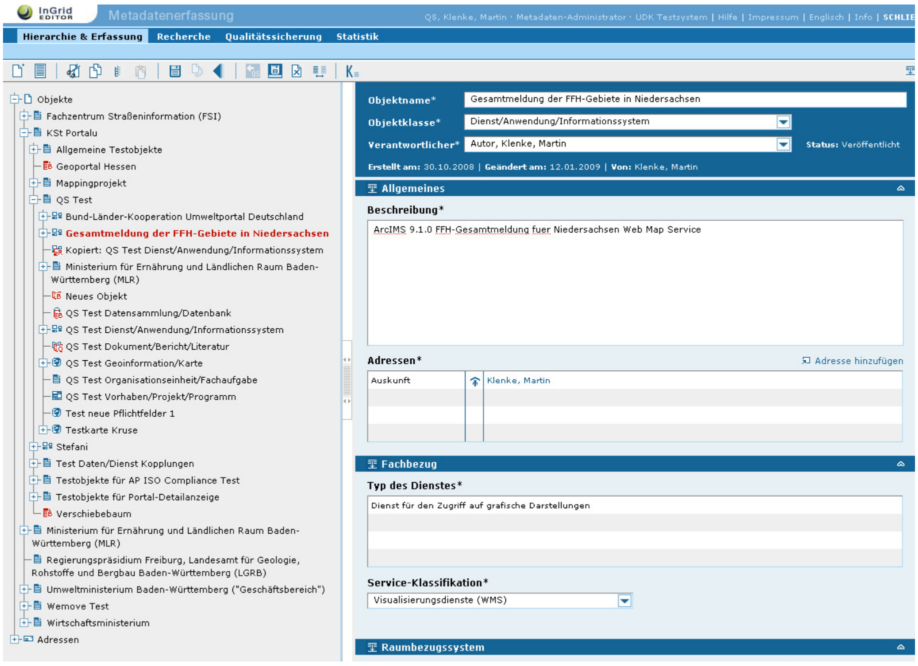


Figure 3: InGrid®Editor

PortalU® is optimized for discovery of geographical and non-geographical data by searching after the describing metadata. For example it is possible to visualise a map in the PortalU® map viewer by one click on a link in the related metadata. Also it is possible to search for example for time series of monitoring data via an internal interface in the original database and visualise the data by use of the original visualisation methods of the database in the web.

A user friendly Editor, the InGrid®Editor (Figure 3), for collection and maintenance of metadata is integrated as a core component of the InGrid®Catalog. It includes some highlights like hierarchically organized metadata for the quick and user oriented access to the metadata objects or separated addresses to minimize time and effort to maintain them. There are different wizards to add the metadata for services, to add information about documents or to generate a list of keyword for a metadata set. Furthermore a workflow management guaranties high quality of the metadata. So the responsible user will be informed about expired metadata sets and a quality manager can control the content of the new or changed metadata sets.

5. TOWARDS SEIS

It is planed to expand PortalU® to fulfil reporting obligations from the regional towards the national authorities as well as the reporting obligations from the federal government to the European commission. For this purpose the direct access to monitoring data shall be integrated into the system. Altogether a basis shall be established to visualize the data and to create reports in form of standardized documents, tables and spreadsheets. In near future a technical concept for these purposes will be developed.

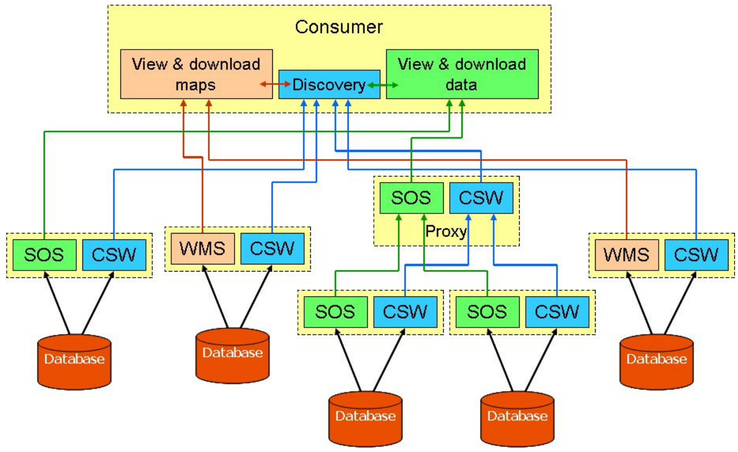


Figure 4: Network of SOS-Services

Today the software InGrid®, on which PortalU® is based, integrates standardised interfaces for catalog services (CSW) and for mapping services (WMS). In future the interfaces shall be advanced to support sensor observation services (SOS) for transfer and visualization of monitoring data (Figure 4). The metadata will be automatically generated by using the GetCapabilities and the DescribeSensor operations. It will be indexed and stored in the InGrid®-like way. The metadata can be discovered by the InGrid® portal surface or by the CSW-interface (Figure 2). The exchange of monitoring data will be routed via the SOS-interface. The portal surface of InGrid® will be extended by a SOS-viewer to visualize monitoring data.

6. CONCLUSIONS

Organisationally PortalU® acts on different administrative levels. It integrates the access to a large number of heterogeneous and geographically distributed information of many providers from different administrative levels.

Technically it is a distributed system that leaves the information at the providers but offers a single access point to all information to the users. It supports all interfaces that are in discussion in the scope of environmental or geographical information. It complies with all requirements of the INSPIRE directive and with ISO or OGC standards related to geographical information.

So from an organisational as well as from a technical point of view PortalU® could be a prototype for a pan-European shared environmental information system.

REFERENCES

Vögele, T., F. Kruse, and O. Karschnik, The gein® 2.0 Information Broker for Environmental and Geospatial Data, *Proceedings of the 10th EC GI & GIS workshop*, Warsaw, 2004.

Vögele, T., M. Klenke, F. Kruse and S. Groschupf, A New and Flexible Architecture for the German Environmental Information Network, *Proceedings of the 19th International Conference on Informatics for Environmental Protection*, Brno, 2005

Vögele, T., M. Klenke, F. Kruse, H. Lehmann, T. Riegel, Easy Access to Environmental Information with PortalU, *Proceedings of the 20th International Conference on Informatics for Environmental Protection*, Graz, 2006

Klenke, M., F. Kruse, H. Lehmann, T. Riegel and T. Vögele, InGrid 1.0 – The Nuts and Bolts of PortalU, *Proceedings of the 20th International Conference on Informatics for Environmental Protection*, Graz, 2006

Vögele, T., M. Klenke and F. Kruse, Metadata Creation and Management of Distributed Data Catalogs with PortalU® and InGrid® 1.1, *Proceedings of the 21th International Conference on Informatics for Environmental Protection*, Warsaw, 2007

Vögele, T., M. Klenke, F. Kruse, H. Lehmann, C. Giffei, PortalU® – a Tool to Support the Implementation of the European Environmental Information Directive, *Proceedings of the 21th International Conference on Informatics for Environmental Protection*, Warsaw, 2007

Klenke, M., F. Kruse, S. Uhrich, C. Giffei, S. Peters, Shaping PortalU®/InGrid® to meet the INSPIRE requirements - First experiences and some considerations on how to efficiently organize the German spatial data infrastructure as INSPIRE node, *Proceedings of the 22th International Conference on Informatics for Environmental Protection*, Lüneburg, 2008

Uhrich, S., F. Kruse, M. Klenke, C. Giffei, S. Peters, The information technology of the German Environmental In-formation Portal PortalU® as tool for the European Level?, *Proceedings of the 22th International Conference on Informatics for Environmental Protection*, Lüneburg, 2008

European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

SOSI
Spatial Observation Services & Infrastructure

G. Triebnig[b], L. Kucera[c], S. Meissl[b], M. Novacek[a], W. Weghofer[d], L. Zsidai[e]

[a] *ANF Data, Pujmanové 1221/4, 140 00 Prague 4, Czech Republic (milan.novacek@siemens.com)*

[b] *EOX IT Services, Thurngasse 8/4, 1090 Vienna, Austria (office@eox.at)*

[c] *GISAT, Charkovská 7, 101 00 Prague 10, Czech Republic (lubos.kucera@gisat.cz)*

[d] *Siemens Austria, Siemensstrasse 88-92, 1210 Vienna, Austria (wolfgang.weghofer@siemens.com)*

[e] *Siemens Hungary, Gizella út 51-57, 1143 Budapest, Hungary (laszlo.zsidai@siemens.com)*

Abstract: SOSI is a project for development and verification of innovative infrastructure and services within the context of land monitoring and Earth observation initiatives at European and Member State levels. The information managed within SOSI includes official land cover products and satellite observation data. The project's results contribute to the concept definition of the Shared Environmental Information System (SEIS) emphasizing symmetrical online sharing of environmental information as opposed to unidirectional environmental reporting. In a pre-operational set-up, involving the European Environment Agency (EEA), four European Member State sites, and the European Space Agency (ESA), a decentralized information system is demonstrated integrating distributed data and processing services as well as interactive geoportal access via multi-lingual graphical user interfaces to geospatial information products provided at multiple levels and content granularities. Practical and more so marketable solutions are investigated within SOSI respecting the INSPIRE directive, guidelines and implementation rules for spatial data service models, engineering and technology. The paper presents the architectural and technological approach along the five ISO/RM-ODP viewpoints and concludes with the main recommendation that service provisioning in the SEIS framework can strongly benefit from a cost and time saving (re-)utilization of the Service Support Environment (SSE) technology provided by ESA and from building upon the related rich experience.

***Keywords**:* Environment monitoring, satellite Earth observation, land cover, Service Oriented Architecture.

1. BACKGROUND AND CONTEXT

1.1 SEIS – a European Perspective

According to the European Union's 6th Environmental Action Programme aiming at a Sustainable Development Strategy integrated assessments of environmental information are becoming the trend. This nurtures the vision of a Shared Environmental Information System (SEIS) which's scope is to improve, modernise, streamline, and connect environmental information systems in Europe and world-wide (see Ref. EC, 2008). From the European Environment Agency's (EEA's) perspective (Steenmans, 2009) SEIS is about:

- Sharing (organisation): Political commitment (legislation); Partnership (win-win); Networking (connecting);

- Environmental Information (content): Horizontal integration (data centres); Vertical integration (local to global); Online access - real time for policy makers and public;
- System (infrastructure and services): Existing ICT Infrastructure; INSPIRE, Reportnet, GMES, etc.; New e-Services (e-Government).

In this list the SOSI project addresses primarily the items “networking”, “online access” and all items under “System”, i.e. Europe's infrastructure backbone for environmental services.

1.2 SOSI – Verification of Infrastructure and Services Concepts

SOSI is a development of innovative “Spatial Observation Services and Infrastructure” within the context of Land Cover (LC) monitoring and Earth Observation (EO) initiatives at European and Member State (MS) levels. The project is motivated by the need of verification of today’s SEIS infrastructure and services concepts (see Section 1.1). The project’s objective is to demonstrate distributed data and processing services as well as decentralized access at multiple levels, languages and content granularities. SOSI is a:

- Project (time/resources limited) contracted by the European Space Agency (ESA);
- Service-Oriented Architecture (SOA) concept;
- Demonstrator system (operated over a limited time period);
- Concept verification activity performed jointly with EEA, MS, and ESA.

The verification results and related recommendations are the most important SOSI project deliverables.

1.3 Coherence with INSPIRE, GEOSS, and GMES Architectures

A key goal of the SOSI project is to ensure coherence with architectures dealt with by the programmes, projects, and standards initiatives given in Table 1.

Table 1. Sources of Standards Requirements for SOSI.

Programme/Project	Explanation (Relevance for SOSI)
INSPIRE	Directive 2007/2/EC of the European Parliament and Council establishing the Infrastructure for Spatial Information in the European Community. Normative implementing rules and informative guidance documents on the aspects of metadata, interoperability, network services, data and service sharing are or will be adopted.
GEOSS	Global Earth Observation System of Systems, intergovernmental programme; core architecture in alignment with OpenGIS Consortium (metadata/catalogue) standards.
GMES - HMA	European Global Monitoring for Environment and Security initiative; including Harmonised Access to heterogeneous EO satellite Missions.
OGC	OpenGIS Consortium leading the development of standards for geospatial and location based services.
GIGAS	Support Action to the European Commission (EC) for identification and definition of needs to enable a full integration of INSPIRE, GMES, and GEOSS architectures.
ISO/TC 211	Technical Committee under the International Organization for Standardization responsible for digital geographic information; issues International Standards, Technical Specifications and Technical Reports (encoding, metadata, services, etc.).
CEN/TC 287	TC under the Comité Européen de Normalisation for digital geographic information in Europe; close cooperation with ISO/TC 211 to avoid duplication.
LMCS/Geoland	GMES Land Monitoring Core Service, development of value-adding (mapping and information) services in the EC Framework Programme 7 project geoland2.

In spite of the complexity and dynamics of the listed programmes and projects the SOSI project can build upon the latest status and specifications through formal involvements of SOSI project team members within OGC, HMA/DAIL, GIGAS, and geoland2.

1.4 Service Support Environment (SSE)

SOSI utilizes the Service Support Environment (SSE) (ESA, 2008) which is a freely available overarching infrastructure permanently hosted and operated by ESA that facilitates the set up of operational services for EO. Currently more than 290 services are connected to SSE serving over 1700 registered professional & science users. SSE is beneficial for Service Providers (SPs):

- Allowing SPs to fully control their own infrastructure;
- Reducing service set-up/demo costs and time to market;
- Allowing an easy re-use of generic or basic services;
- Allowing SPs to concentrate on their own specific skills and added value;
- Combining and chaining service elements within the supply chain;
- Enabling the cooperation between SPs, etc.

1.5 Demonstration System

The SOSI Demonstration System under implementation is shown in Figure 1. Four servers hosting MS-level services are installed in Austria, Czech Republic, Hungary, and Luxembourg. A European-level service is provided by the EEA. The SSE hosted at ESA provides the geoportal, workflow, and registry functionality. The services and their characteristics are further explained in Chapter 4. They are "consumed" by various users on the Internet utilizing different kinds of clients.

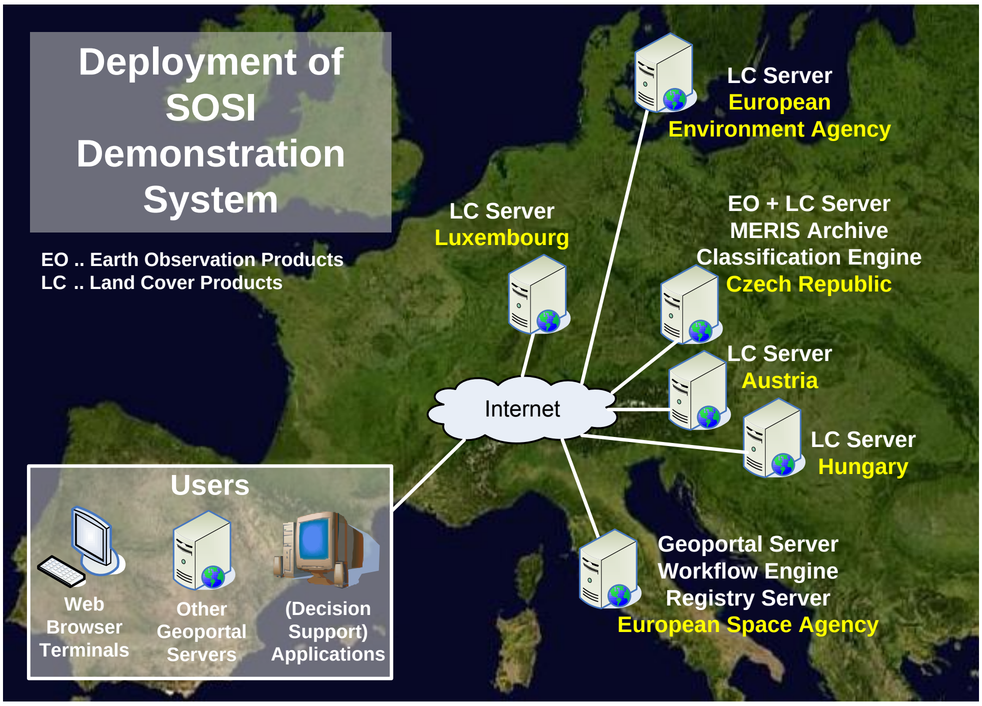


Figure 1. SOSI Servers in Member States and at European Agencies.

1.6 SOSI Architecture Viewpoints

The Reference Model of Open Distributed Processing (RM-ODP, ISO/IEC 10746-1,2,3,4) is an architecting standard. The model is widely adopted as best practice approach by the geospatial data system engineers. Presentations of the five RM-ODP design viewpoints, tailored to SOSI, are the subjects of the following Chapters.

2. ENTERPRISE VIEWPOINT

2.1 From Reporting to Sharing

The SOSI design concept is driven by SEIS principles (Steenmans, 2009) that data and information shall be:

- Managed as close as possible to its source;
- Collected once and shared with others for many purposes;
- Readily available to public authorities to easily fulfil reporting obligations;
- Accessible to enable end-users to make comparisons at appropriate geographic scale;
- Available to the general public after due consideration of the appropriate level of aggregation, given possible confidentiality constraints, and at national level in the national language(s).

2.2 Eionet and Evolution of Reportnet

SOSI receives design guidelines and standards (e.g. the GEMET Thesaurus) established by EEA and within Eionet which is a partnership network of EEA and its member and cooperating countries involving approximately 900 experts and more than 300 national institutions. The network supports the collection and organization of data and the development and dissemination of information concerning Europe's environment.

The SOSI implementation is targeted to be linked with Reportnet which is Eionet's infrastructure for supporting and improving data and information flows. Reportnet is based on a set of inter-related tools and processes which all build on the active use of the World Wide Web. SOSI is specifically associated with the objectives and functionalities of the EEA Land Use site which offers live maps and EEA data centre services under the address http://www.eea.europa.eu/themes/landuse. It is required, for example, that SOSI network services can be accessed via the existing EEA Land Use site.

2.3 GMES Fast Track Service on Land Monitoring

Essential innovative functionality is provided by SOSI for orchestration of network services within workflows. Product generation chains which, in the past, transferred data between processing nodes on offline physical media shall be demonstrated in SOSI to be handled, in the future, via online management and data dissemination procedures. The GMES Fast Track Service on Land Monitoring which was implemented by EEA, EC and ESA is considered as template for a SOSI demonstration example. It delivered the following products:

1. Ortho-rectified satellite images for the reference year 2006 (+/- 1 year);
2. European mosaic based on ortho-rectified satellite imagery (IMAGE2006);
3. Corine Land Cover changes 2000-2006;
4. Corine Land Cover map 2006 (CLC2006);
5. High resolution core Land Cover data for built-up areas, including degree of soil sealing, 2006;
6. High resolution core Land Cover data for forest areas 2006.

For the CLC2006 and high resolution built-up areas products open access is granted and free dissemination data policy is applied.

The related requirements established for the SOSI Demonstration System are to:

- Mimic the current geoportal access to CLC2000, CLC2006 and built-up area products keeping the products close to their geographically distributed generation services, i.e. not, as currently done in a centralized product repository;
- Demonstrate a processing chain fully managed as an online workflow spanning these distributed product generation services.

2.4 Specific Requirements in Member States

The requirement for the SOSI processing workflow demonstration described in section 2.3 is instantiated as a semi-automated land cover product generation chain sourced by an Envisat MERIS satellite EO product archive in the Czech Republic. As prerequisite for this demonstration implementation a user service facility is implemented at the MERIS archive supporting data discovery and downloading services. A case study focusing on benchmarking of MEEO Soil Mapper classification solution is another SOSI activity run in the Czech Republic.

Beside the Czech MERIS EO services a national Czech SOSI LC geoportal is to be implemented. Except for the European LC datasets described above it will also include selected national GMES LC datasets and samples of low resolution LC products based on MERIS imagery. A SOSI geoportal for LC data access is also set-up in Hungary which has to support a Hungarian national language user interface.

3. INFORMATION VIEWPOINT

3.1 Metadata Models

Service and Collection Metadata are used in SOSI to provide the details for machine-to-machine communication (operations, coupling of operations) but also contain descriptive information targeted at human readers (identifying properties of the described data or services). These metadata are in XML format, their standard references are: ISO 19115, ISO 19119, and OASIS UDDI v2 Metadata.

EO Product Metadata are specified in XML format and based on Geography Markup Language (GML) Encoding Specification, OGC 03-105 and 02-069 (ISO 19136), and GML 3.1.1 Application Schema for EO Products, OGC 06-080.

3.2 Spatial Data Sets

Various delivery formats are expected by the community (e.g. GeoTIFF, JPEG2000, HDF). Conceptually (as in SSE and HMA), spatial data sets addressed in SOSI can be classified:

- With respect to production mechanism as Stream (regularly disseminated with pre-defined frequency, with given time window and given area); Cumulative (disseminated at package completion); Standalone (one-off dissemination after single request);
- With respect to triggering as Pre-defined Data Sets (well defined user needs/area defined in advance); or On-demand Data Sets (specification that cannot be identified in advance, e.g. emergency monitoring).

The SOSI Demonstrator implements the above cases Standalone, Pre-defined, and On-demand Data Sets when providing access to the following spatial data sets:

- GMES Fast Track high resolution core land cover data for built-up areas, including degree of soil sealing 2006; European-level seamless and MS-level products; updates are foreseen generated from 2009 source EO data acquisitions;
- Corine Land Cover - CLC2000 and CLC2006, and specific national LC products;
- Envisat MERIS data and derived LC Products (currently under specification).

4. SERVICE VIEWPOINT

The informative documents “INSPIRE Technical Architecture Overview” (JRC, 2007) and “INSPIRE Network Services Architecture V3.0” (JRC, 2008) are key references for the SOSI service design. The SOSI architecture (conceptual and Demonstrator System) includes the services listed in Table 2. The SOSI Service Metamodel is shown in Figure 2. Indications are given about which service is to be supported by SSE Technology – see Chapter 6. Transactional functions for ingesting the metadata, spatial data sets, and register data into the indicated data sources are considered to be parts of the inherent data base management functions, though a use of network services could be considered (e.g. based on OGC WFS-T or WCS-T technology).

Table 2. SOSI Services.

GEOSS Taxonomy	**SOSI Services** (INSPIRE service naming used, as applicable)
Geographic human interaction services	SOSI geoportals based on the SSE graphical user interface
Geographic model/ information management services	SOSI spatial data discovery service
	SOSI spatial data view services
	SOSI spatial data download services
	SOSI invoke spatial data services (value adding service chains): • For MERIS satellite data spatial and thematic processing • For LC statistics processing (Built-up areas aggregated according to Corine classes)
Geographic processing services – spatial	MERIS satellite data ortho-rectification service
Geographic processing services – thematic	MERIS satellite LC classification service (MEEO solutions)
	Built-up areas aggregated according to Corine classes (GIS service demonstration example)
Geographic system management services	SOSI services discovery service
	SSE services registry service
	SSE rights management
Geographic processing services – temporal	None
Geographic processing services – metadata	
Geographic communication services	

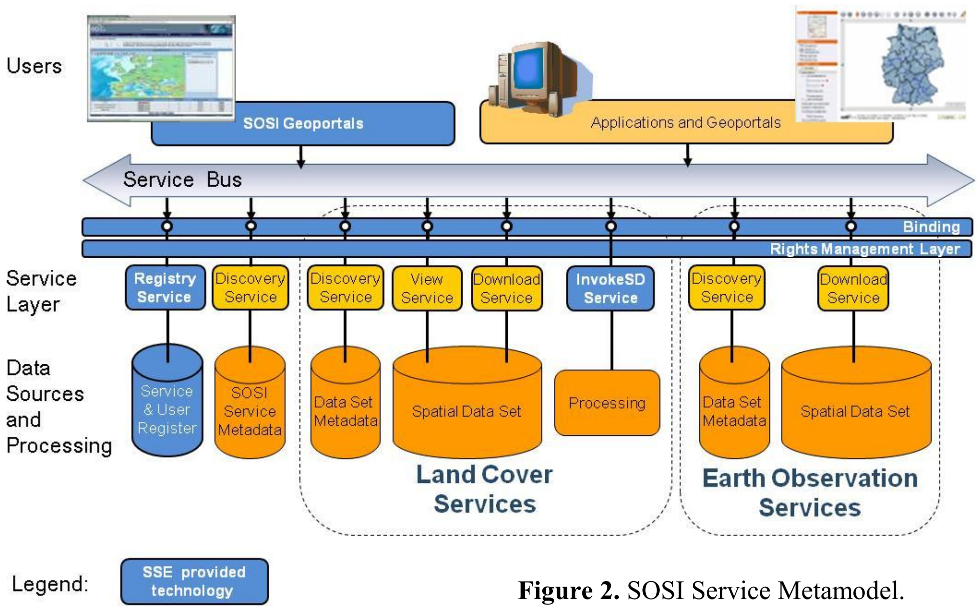


Figure 2. SOSI Service Metamodel.

5. ENGINEERING VIEWPOINT

SOSI is designed as loosely-coupled system based on networked services enabling distribution through functional decomposition into objects which interact at standard interfaces. Conceptually, SOSI aims to be an open system into which additional services can be added, without need for additional software development, assuming interfaces are precisely followed.

SOSI is designed scalable to cope with growing user numbers. This means that it is possible, if needed, to add additional CPUs and balance the load on multiple machines, in a way that is transparent to the SOSI application software. In the SOSI Demonstration System set-up this is demonstrated for the SOSI geoportal hosted on the computer centre facilities at ESA distributing the load of all SSE induced traffic to a number of CPUs.

SOSI places no restriction on the granularity of a (Web) service that can be integrated. The grain size can range from small (e.g. a component that must be combined with others to create a complete business process), to large (for example an entire value chain).

6. TECHNOLOGY VIEWPOINT

SOSI is based on Service Oriented Architecture (SOA) principles. The drivers of the architecture are open standard specifications from the bodies ISO, OASIS, W3C, and OGC. Figure 3 shows the software which is used in the SOSI Demonstration System. The discovery, view, and downloading services are implemented entirely on Open Source tools implementing the shown Web server protocol standards. Open Source software is also used for back-end database management.

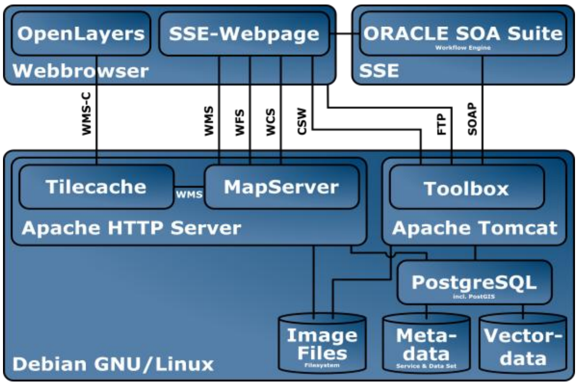


Figure 3. SOSI Software and Web Protocols

The SSE components which are deployed as platforms in the SOSI project are:

- SSE Portal Server, provides the SOSI workflow engine (ORACLE SOA Suite 10.1.2) which orchestrates distributed services; several wizards for service registration at the SSE are available;
- The SSE WebMapViewer forms the cartographic display of the SOSI geoportal. It provides the functionality to define Area(s) Of Interest (AOI) by end-users as input to services and for receiving portrayal and download service results (see Figure 4);
- The SSE Toolbox provides a wrapper from legacy services at SOSI service providers to SOAP web services. It supports synchronous and asynchronous service calls and includes a FTP server. The runtime environment provides a rich stand-alone test centre for services. There is also a development environment providing an Eclipse plug-in for the whole service life-cycle (i.e. development, testing, debugging, configuration control, deployment, etc.).

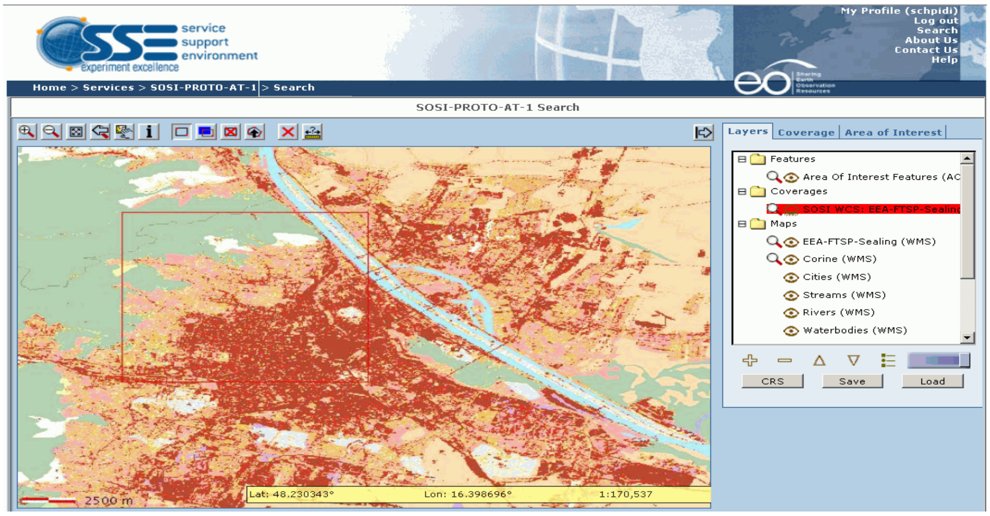


Figure 4. SOSI Geoportal Supporting Product & Service Discovery, Versatile Area of Interest Definition, WMS-based Viewing, and WCS-based Dowloading.

7. CONCLUSIONS AND RECOMMENDATIONS

In the search for suitable approaches to modernise and simplify the collection, sharing and use of the data and information required for the design and implementation of environmental policy in Europe, the SOSI project is contributing technical concepts for strategically important land cover (LC) services and infrastructure components. The full demonstrations of SOSI LC service functionalities will be available in June 2009, the Earth Observation service and LC generation service chaining demonstrations will be ready for verification towards the end of 2009. The conceptual designs of SOSI and first prototypical system implementations are, however, available and the following conclusions and recommendation can therefore be made.

SSE, which is based on substantial experience with numerous operational services and users, should be considered as a valid and capable technological platform by those who need to implement services in the SEIS framework quickly, cost-efficiently and based on latest knowledge. A wide range of international architecture and technology standards is already pre-digested in the SSE tool suite helping service providers and their engineers to get up quickly on the learning curve.

From the SOSI perspective the following enhancements of SSE are required:

- A federated user management and security layer needs to be implemented in the SSE Portal Server (e.g. in the service registration wizards) as well as in the SSE Toolbox;
- In order to achieve further INSPIRE conformity the SSE WebMapViewer needs upgrading with respect to SOAP binding for OGC Web services (currently only Key-Value-Pair binding supported);
- A modernization of WebMapViewer map interface needs to incorporate latest technologies such as AJAX. Also required are enhancements of client functionality for optimum exploitation of INSPIRE Network Services (e.g. automated bridging process from view to download).

It is recommended to further study, and potentially adapt and deploy the SOSI service chaining mechanism for the next planned Europe-wide LC product generation based on 2009 EO data. An implementation of a fully online work flow appears to be feasible.

Aligned with the SOSI approach, EEA and ESA should consider a ramping up of activities towards conducting an official OGC Pilot, inviting a larger community of MS representatives and user organisations into a collaborative effort that applies open standards for interoperability to achieve user objectives in a surrounding which is representative for operational use. With SOSI results at hand further consensus on architectures for geospatial information systems should be aimed at and carried forward into sustained operations.

ACKNOWLEDGMENTS

The project is conducted under contract of the European Space Agency ESA (ESRIN Contr. No. 21776/08/IL-G and ESTEC PECS Arrangements with CZ and H) and the strong in-kind steering support by the European Environment Agency EEA. The authors wish to thank Mr. S. D'Elia and Mr. P.G. Marchetti from ESA for fostering and supporting this initiative, Mr. A. della Vecchia for the technical cooperation as well as Mr. C. Steenmans, Mrs. A. Sousa, Mr. J. Bliki, and the LUDC team from EEA for essentially having formed the project idea, and for accompanying its execution. The preparedness by the company Geoville to host the Luxembourg SOSI instantiation is very much appreciated. The contributions of the company Spacebel regarding SOA and SSE knowhow and of the company MEEO regarding automated LC generation are very valuable to the SOSI project.

REFERENCES

EC COM(2008) 46 final - Towards a Shared Environmental Information System (SEIS).

ESA – European Space Agency, Service Support Environment, 2008, http://services.eoportal.org.

JRC – Joint Research Centres, INSPIRE Technical Architecture – Overview, V1.2, 2007, http://inspire.jrc.ec.europa.eu/reports.cfm.

JRC – Joint Research Centres, INSPIRE Network Services Architecture, V3.0, 2008, http://inspire.jrc.ec.europa.eu/reports.cfm.

Steenmans, C., The Future of Environmental Information, presentation at workshop "*Towards Shared Environmental Information in Europe: Opportunities and Challenges for Integrated Research",* Centre for Ecology & Hydrology, Wallingford, UK, 28-29 January 2009*,* http://www.ceh.ac.uk/SEIS/documents/EEASEISCEH.pdf.

Sponsors

HTS-Brno

Honeywell

Integrated Design Center

HTS-Brno

At Honeywell, our engineers are building a world that is safer and more secure…more comfortable and energy efficient…more innovative and productive. We are Honeywell.

- World-class engineering to drive business growth
- Strategically-linked into Honeywell businesses
- Technology cross-over between engineering domains
- Engineering excellence for quality, speed and productivity
- Functional excellence enabling mission dilivery

Your career in the Honeywell Brno Design Center… The most talented people. World-class technologies. The best of both worlds!

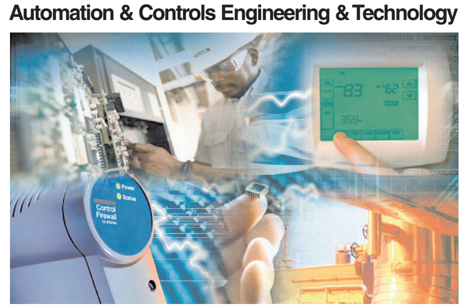
Today there are over 2,700 Honeywell employees working in the Czech Republic engaged with aerospace, automation and transportation businesses in Brno, Prague and Olomouc. And underscoring its strong commitment to invest in the Czech Republic and to enhance our technology presence here, Honeywell continues to expand its 650-person design center operations in Brno.

Founded in 2003, the Integrated Design Center in Brno is part of Corporate Honeywell Technology Solutions which has research centers in U.S, China, India and the Czech Republic. Robert Aasen, Director of the design center explains how the role of the center is growing in importance. "Engineers with disciplines in software, electronics, mechanical, control systems, turbomachinery, and guidance and navigation undertake development for three of Honeywell's business groups – Aerospace, Automation and Control Solutions, and Transportation Systems. Development work covers a wide range of products and solutions such as flight control systems, homes and buildings control and regulation units, solutions for security systems and turbochargers for commercial and passenger vehicles."

Aerospace Engineering & Technology

- Guidance and navigation
- Flight control systems
- FADEC (Full Authority Digital Engine Control)
- Engineering test services
- Advanced technology

Automation & Controls Engineering & Technology



- Combustion controls and valves
- Building automation systems
- Home comfort controls
- Water and hydronic heating
- Access and security systems
- Enabling technologies

Transportation Engineering & Technology

- Innovation, variable geometry turbochargers, aerodynamic & vibro-accoustics analysis, bearing systems for:
- Commercial vehicles
- Passenger vehicles, light trucks

Find out more

Employment Contacts
Zuzana Kvetonova, IDC
+420545501357
Jarmila Uherova, Aero
+420532115016
Frantisek Snobl, Turbo
+420545501336
Pavlina Kotrousova, ACS
+420545501342
www.honeywell.cz
www.honeywell.com/careers

Honeywell Brno
Integrated Design Center
Honeywell Technology Solutions
Turanka 96/1236
627 00 Brno/Slatina, Czech Republic
+420-545-501-702
www.honeywell.com

Honeywell

N61-0824-000-000
May 2008

Energy Savings Performance Contracts

Energy Savings Performance Contracts (ESPCs) allow government agencies, schools, hospitals and the like to pay for critical building improvements with energy and operational savings. Under an ESPC, Honeywell identifies areas where it can reduce energy consumption and expenses. The customer pays no upfront costs for the projects selected. Additionally, the savings the upgrades generate fund the work over the length of the contract, often spanning several years. Honeywell guarantees the savings so there is no risk.

Application: Fort Bragg

Executive Order 13123 calls for federal agencies to reduce their energy consumption 35 percent by 2010. With that mandate in mind, Fort Bragg - the U.S. Army post in Fayetteville, N.C. - began a partnership with Honeywell to develop a strategy to meet this target.

Leadership at the post also wanted to decrease emissions, repair, modernize and improve facilities, increase energy security, and improve the overall quality of life at Fort Bragg. Honeywell initially implemented a variety of basic demand-reduction measures, including high-efficiency lighting retrofits, control system installation and enhancements, HVAC unit upgrades and replacement and building envelope modifications.

The savings from these projects were significant, but not on a magnitude to accomplish all of the goals. So Honeywell and the Army developed a method to capture energy supply savings via an ESPC. In order to manage the risks associated with complicated supply contracts, it becomes necessary to integrate energy demand and energy supply management through information technology and control systems. This shifted the focus from energy savings to energy management - a.k.a., active demand management. As a result, Honeywell was able to negotiate a new power supply tariff based on real-time power pricing, negotiate new natural gas contracts and improve natural gas usage and build a 5-megawatt cooling, heating and power generation system designed to improve energy infrastructure and security.

We are building a world that's safer and more secure…
More comfortable and energy efficient…
More innovative and productive…
We are Honeywell

Market Trends and Opportunities

There is compelling evidence to suggest that traditional demand-reduction improvements - lighting upgrades, for example - have limited scope in the face of today's volatility energy market. As a result, more organizations are turning to the type of integrated program that Fort Bragg employs. With this approach, the whole is greater than the sum of the parts.

Honeywell is also adding more renewable and alternative energy sources to ESPC projects. This includes clean coal, wind power, photovoltaics, biomass, geothermal and other green technologies.

ESPC Features:

- Pays for facility upgrades with energy and operational savings
- Places the burden of performance and cost savings on Honeywell
- Offers an alternative to the traditional bid and spec process

Customer Benefits

- Funds critical projects that would be postponed or neglected
- Increases comfort and energy efficiency, and combats rising energy costs
- Helps address compliance issues by decreasing consumption, emissions and reliance on traditional energy sources
- Doesn't impact operating costs or require additional taxpayer dollars

Since 1992, Honeywell has completed more than 4,000 energy-saving projects and delivered $3.1 billion in guaranteed energy and operational savings.

Find out more

For more information on Honeywell Energy Savings Performance Contracts, please contact Aaron Parker at 952-656-1721 or aaron.parker@honeywell.com

Honeywell
Automation and Control Solutions
1600 Utica Avenue South, Suite 300
St. Louis Park, MN 55416
www.honeywell.com

N61-0325-015-000
August 2005

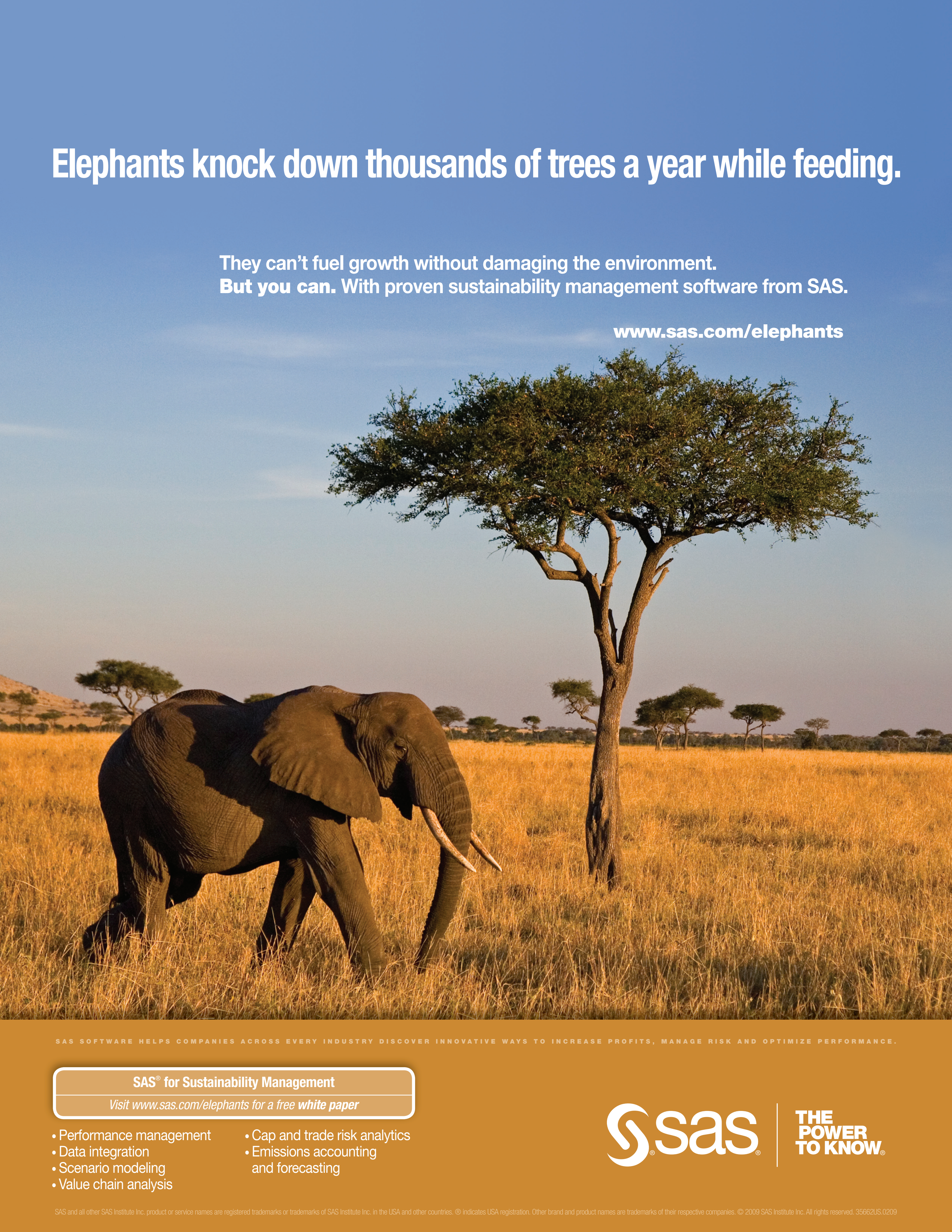
Honeywell


www.praha.eu
PRAHA.EU
PRA HA
PRA GUE
PRA GA
PRA G
Your Gateway to Prague
Portál hlavního města Prahy - Windows Internet Explorer
http://www.praha.eu/jnp/cz/home/index.html
Virtual Prague Tours
Transport
Cycling in Prague
Exchange Rates
Help
Portal of Prague
Today -4°C
Tomorrow -1°C
Next Day 0°C
Search
Residents
Business
Visitors
Extra
Opencard
Introduction
City Hall
Residents' Service
Useful Contacts
Historic Prague
Religion
Health Care
People with Disabilities
Transport
Culture
Social Sphere
Gay and Lesbian
Environment
Air Quality
Water Quality
Waste
Noise
Through Expat's Eyes
City
Public Safety
Prague and EU
Home Page > Residents > Environment
Environment
Quality of the environment is of primary importance to all of us.
The city of Prague, its council and other city authorities ascribe great importance to the protection and care of the environment in the city as well as to providing quality data in order to make decisions and to inform the public. We believe you will find all the information you need. At the same time we welcome all comments and suggestions concerning these new pages.
Envis 4
The objective of ENVIS 4 project is to create a system that would provide selected municipal districts with enhanced accessibility and quality of information on the environment. The information will mostly relate to the air quality, roadway noise levels, landscape and verdure. The project makes use of different information and communication technologies including geographical information systems, databases, modelling, evaluation of aerial photographs and internet information services.
News
Černý Most is Going to Turn Green
A 170-hectare park twice as large as Stromovka, will be built between Černý Most, Dolní Počernice and Hostavice in four years time.
more
We Will Have a "Forest" in the City
Prague planners have selected some forests which will be very minimally managed in the future. This concerns 65 forests spreading over an area of 107 hectares.
more
Updated: November 5th, 2008
Generate PDF
Print
Send
Contacts
Webmasters Contacts
Privacy Statement
Site Map
© 2006-2009 Magistrát hl. m. Prahy
I Know How I Drive
Are you interested in the traffic offences you are suspected of committing in Prague?
Cameras
Do you want to get up-to-date information about the traffic situation in Prague?
InterSpis
You can communicate with the municipal offices from the comfort of your own home and follow the progress of your application on-line
Prague, the City of a Hundred Spires
Experience a virtual tour of the city; see views from the most famous, as well as the lesser known of Prague's unique towers.
Events ■ Information ■ E-Office ■ Telephone Centre

Improve Business Performance

ARIS: Nástroje a konzultační služby pro řízení procesů

www.ids-scheer.cz


Elephants knock down thousands of trees a year while feeding.
They can’t fuel growth without damaging the environment.
But you can. With proven sustainability management software from SAS.
www.sas.com/elephants
SAS SOFTWARE HELPS COMPANIES ACROSS EVERY INDUSTRY DISCOVER INNOVATIVE WAYS TO INCREASE PROFITS, MANAGE RISK AND OPTIMIZE PERFORMANCE.
SAS® for Sustainability Management
Visit www.sas.com/elephants for a free white paper
• Performance management
• Data integration
• Scenario modeling
• Value chain analysis
• Cap and trade risk analytics
• Emissions accounting and forecasting
THE POWER TO KNOW®
SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies. © 2009 SAS Institute Inc. All rights reserved. 35662US.0209

GS1 and European Union regulation in Waste Management

7. What advantages does the usage of the GS1 Global Location Number (GLN) have for the obligor of the battery directive (2006/66/EC)?

Article 17 states that each producer has to be registered. The GS1 System enables a producer to use the same identification in every country. Registration procedures are simple and can be duplicated. So it is possible to save costs for Member States and producers.

Advantages in comparison to a national number is that it is unique globally enabling the worldwide interoperability (you need a unique number for databases) is guaranteed. This functionality also enables the RFID using the EPC structures and enables electronic messaging.

8. What are advantages of using the GS1 system in establishing the material identification and information systems for the packaging and packaging waste directive (94/62/EC)?

The Commission shall in accordance with Article 8(2) establish a material identification system related to the nature of the packaging materials used. Article 12 stipulates that the Member States shall take the necessary measures to ensure that databases are established on a harmonised basis to monitor the implementation of the Directive.

By using the GTIN identifier for a product and the central databases, product packaging material attributes (nature and quantity) can be globally recorded for each trade item at primary, secondary and tertiary levels. This will provide the Commission with reliable data about packaging waste material on products. All trading partners across Europe will use the same data. Therefore, it will contribute to enabling EU Member States to supply more detailed and harmonised data waste material.

GS1 in Europe

T +36 1 412 3947
E info@gs1eu.org
www.gs1.eu

GS1 in Europe

QUESTIONS & ANSWERS
on GS1 and Waste Management

About GS1

1. Who is GS1?

GS1 is a neutral not-for-profit global organisation established over 30 years ago. The GS1 Head Office is based in Brussels and has offices (Member Organisations – MO's) in 108 other countries. GS1 was originally established by the manufacturers and retailers in the food and consumer goods sectors to improve supply chain efficiency. Today, the GS1 System is used by over 1.4 million companies worldwide in numerous different sectors. These sectors now include healthcare, transportation and logistics, aeronautics, defence, chemicals, customs, computer hardware and software, local and international regulatory authorities, and many more. The GS1 Systems are used by both the large multinational and by the small medium size enterprise organisations; by the leading brands and by individual craftsmen. The work of GS1 is funded by the user contributions and by the active participation of the users themselves along with a global secretariat.

2. What does GS1 do?

GS1 designs and implements global supply chain standards. The GS1 Standards provide a framework that allows information about products and services, to be exchanged efficiently and securely for the benefit of businesses and the improvement of people's lives, everyday and everywhere.

The GS1 System consists of a portfolio of standards that address how products, assets and services are globally and uniquely identified, how data can be represented in data carriers – bar codes and radio frequency identification tags, and how data can stored and exchange on and between databases using standards data formats and electronic messages.

Today, over 5 billion transactions happen daily based on the GS1 Standards – making them the most widely used standards globally.

GS1 through its 108 MO's support their users with a wide range of services, including training, implementation advice, certification and technical support. GS1 also maintains and updates the standards as user requirements and technology evolves.

3. What relation is between GS1 in Europe and other European associations and organisations.

GS1 in Europe is a collaboration of 45 GS1 Member Organisations from the European Union, EFTA and other countries, with approximately 400 000 user companies.

At both a national and at a European level, the European based GS1 Member Organisations collaborate with several professional, trade, commercial and civil organisations across the different industries. They play an active role in national standard bodies and some are members of the different international standards organisations such as ISO, CEN, UNCEFACT and ETSI .

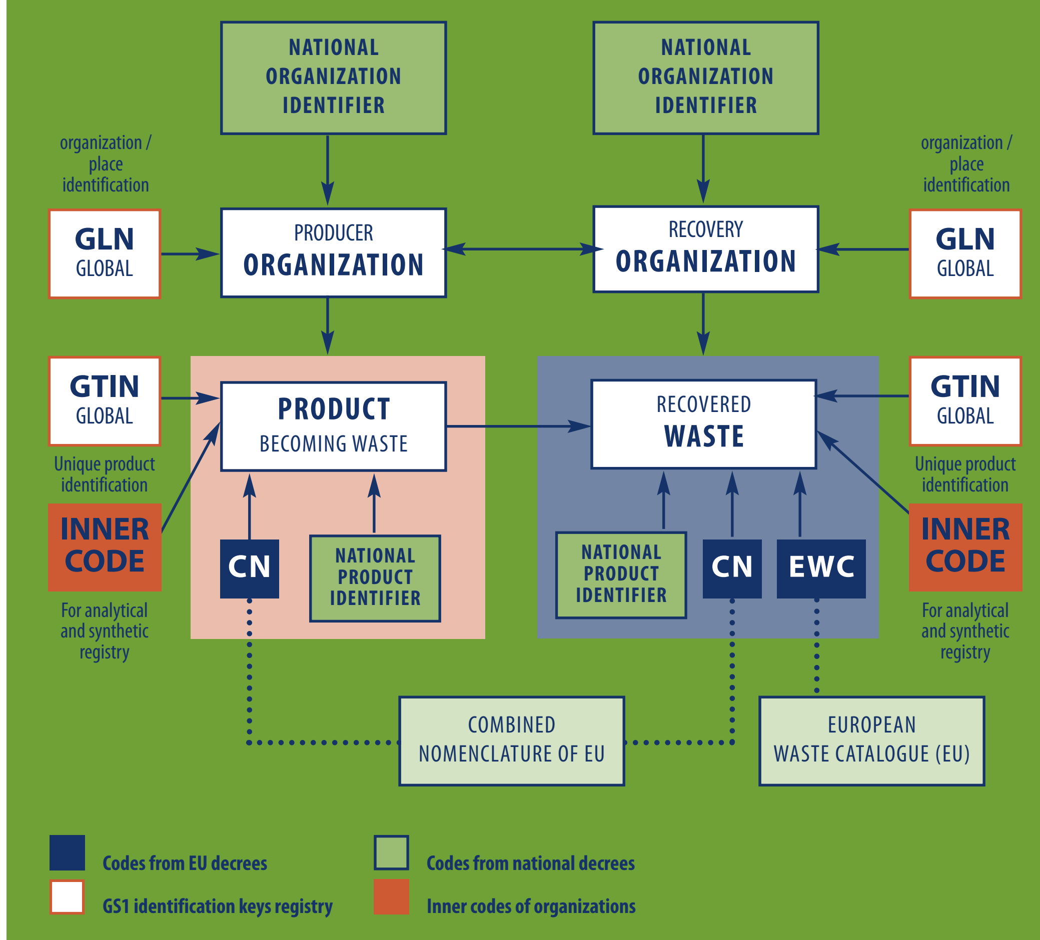
GS1 in Europe is working at establishing a partnership with European Institutions and support the new strategic framework i2010 – European Information Society for Growth and Employment.

GS1 and Waste Management

4. Why is GS1 in Europe involved in waste management?
In industry and trade, the GS1 System is a well known and accepted standard for supply chain management. Using the GS1 System for waste management will complete the product life cycle. The objectives remaining the same for the waste supply chain, i.e., improve efficiency and effectiveness of waste management.
European legislation for products and waste places the responsibility on the "producers" to protect the environment. This affects the 400 000 companies who are already users of the GS1 System in Europe. Therefore, GS1 in Europe launched the Waste Management project to focus on how the GS1 System could be implemented to address the needs of the Regulators and the users. Particular attention was paid to Packaging Waste, Electrical and Electronic Equipment and Batteries and Accumulators.
The objective of the GS1 in Europe Waste Management project is to develop a common strategy on how the GS1 System can facilitate the implementation the EU legislation consistent with systems already implemented by users as they relate to the GS1 System. Also, how the GS1 System can be used by the stakeholders in Waste Management to increase Supply Chain efficiency, accountability and traceability.
The GS1 Systems enables to bridge between commercial supply chain and the environmental/waste management supply chain.

5. How would the GS1 System work in Waste Management?
The GS1 System offers a standard which identifies an individual object globally and uniquely. A product is identified by a Global Trade Item Number (GTIN). Every product has associated attributes – size, dimensions, descriptions, composition, ADR regulations, classification etc. The GS1 System includes a data model which details a structure for storing the data consistently in central databases. Each product is indexed by the globally unique identification GTIN.
Waste material is a product with its own attributes. Waste attribute data might be the descriptions, the classification, the dimensions and weights, specific handling requirements, type of treatment required, material safety data etc. All of the above data is part of the GS1 Data model.
A legal entity, e.g., a company or a physical location, e.g., a warehouse may be identified using Global Location Number (GLN). GLN's, like GTINs, are globally unique and are used extensively in electronic messaging. Each GLN has its associated attributes – names, addresses, company registrations and approvals.
Within the GS1 System an individual logistical unit can be assigned a unique "licence plate" called a Serial Shipping Container code or SSCC. This number globally identifiers the instance of a particular pallet.
Assets can be identified either as fixed (GIAI – Global Individual Asset Identifier) or returnable assets (GRAI – Global Returnable Asset Identifier).
The GTIN, GLN, GIAI, GRAI and SSCC are all registered structures within ISO Standards. They can be incorporated into GS1 bar codes and into RFID tags (EPC Tags).
Therefore, to globally and uniquely identify the waste, the organisations, the shipping units and the handling equipment, GS1 offers solutions that can be implemented in data carriers and

Organization, product and waste identification serves traceability



EXAMPLE OF WASTE SUPPLY CHAIN

referenced in data bases. To exchange information GS1 has implement standards based electronic messaging which allows organisations exchange information using internationally recognised formats.

6. Is the GS1 system compatible with RFID?
In addition to the traditional data carrier, i.e., the barcode GS1 in-conjunction with MIT has adapted and developed Radio Frequency Identification (RFID). Global Standards under the name EPC have now been developed so that RFID can be implemented in open environment for trade. RFID based on EPC is now being implemented across the FMCG sector.
EPC is a solution that will also enable RFID tagging of Waste.
GS1 in Europe is working at establishing a partnership with European Institutions and support the new strategic framework i2010 – European Information Society for Growth and Employment.

Maple™ 12 The Essential Tool for Mathematics and Modeling

12 Redefining Usability

Maple's Smart Document Environment is an intuitive user-interface that automatically captures all of your technical knowledge in an electronic form that seamlessly integrates calculations, explanatory text and math, graphics, images, sound, and more. It has many built-in tools to assist you in analysis and solution development.

12 Extensive Collection of Built-In Mathematical Algorithms

Maple 12 combines the world's most powerful mathematical computation engine with an intuitive user interface. It offers over 4000 mathematical functions, cutting-edge solvers for ODEs, PDEs, and DAEs, highly efficient numeric solvers, and world-leading symbolic solvers.

12 A Vital Component of Your Tool Chain

Maple 12 integrates with CAD, Excel®, and other standard tools. It includes code generation tools for C, Fortran and MATLAB®, and is supported by an extensive range of products for various industrial applications including automotive, aerospace, electronics, energy, and finance.

12 What's New in Maple 12

- A new Dynamic Systems package offers a large selection of analytic and graphing tools for linear time-variant systems, which are essential in control systems development.
- New interactive embedded components include dials, gauges and radio buttons.
- The Exploration Assistant lets you create interactive mini applications to explore the parameters of expressions.
- New plot types include dual axis plots, polar axis plots, and specialized engineering plots.
- Improvements to solvers for ODEs, PDEs and DAEs continue to strengthen Maple's world-leading position in numeric and symbolic differential equation solving.
- The Discrete Transforms collection has been expanded to include numerical wavelet transforms.
- CAD connectivity adds important analysis capabilities to CAD systems.
- MATLAB to Maple code translation allows you to leverage existing work.
- …and more!

Licenses of Maple 12 Professional include the Maple Toolbox for MATLAB, MapleNet, and a 1 year subscription for the Elite Maintenance Program.

To learn more about Maple 12 visit the new Maplesoft Demo Center and see Maple 12 in action!

Go to: www.maplesoft.com/products/maple/demo/

www.maplesoft.com

Maplesoft Mathematics • Modeling • Simulation

MapleSim Roadmap

MapleSim™

High-Performance Multi-Domain Modeling and Simulation

2009

Connectivity Toolbox V1
- Export to Simulink & LabVIEW

MapleSim V2 (April 2009)
- 3D Animation
- Blockset Increase (Hydraulics, Digital Electronics, etc.)
- Advanced Analysis Templates
- Simulation Results Management

Advanced Hydraulics Library

MapleSim Control Toolbox

Tire Component Library

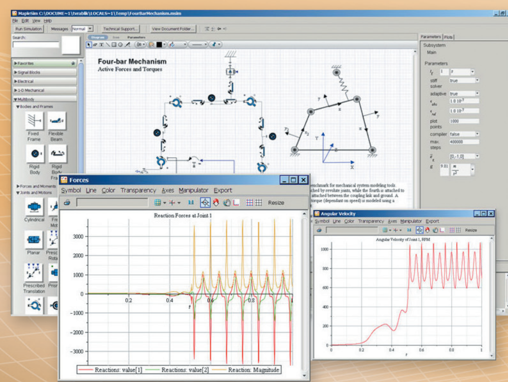
Pneumatic Component Library

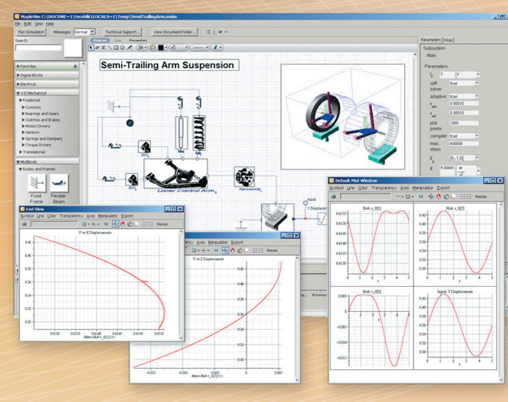
MapleSim V3 (October)
- Interactive Simulations
- Blockset Increase (Advanced Mechanics, Signal Processing, Control Design, etc.)
- Modelica 3.0, Modelica Import

Connectivity Toolbox V2
- Simulink Import

2010

MapleSim V4 (April 2010)
- Blockset Increase (State Machine etc)
- Collision Detection
- 3D Visual Construction
- Events
- Embedded Code Generation





The Elite Maintenance Program (EMP)

provides the best way for you to keep up to date with MapleSim releases. EMP is a renewable subscription that provides members with automatic upgrades to the latest releases of many of Maplesoft's products and solutions. Members of the EMP program will realize significant cost savings over the course of MapleSim's development.

Contact Maplesoft for details.

Maplesoft
Mathematics • Modeling • Simulation
www.maplesoft.com | info@maplesoft.com
Toll-free: (US & Canada) 1-800-267-6583 | Direct:1-519-747-2373

EUMETSAT – Monitoring weather and climate from space

Towards eEnvironment, Prague, Czech Republic, 25-27 March 2009

ABOUT EUMETSAT...

The European Organisation for the Exploitation of Meteorological Satellites is an intergovernmental organisation based in Darmstadt, Germany, currently with 24* Member States (Austria, Belgium, Croatia, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia*, Luxembourg, The Netherlands, Norway, Poland*, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey and the United Kingdom) and six Cooperating States (Bulgaria, Czech Republic, Estonia, Iceland, Lithuania, and Romania).

The main mission of EUMETSAT is to deliver weather and climate-related satellite data, images and products – 24 hours a day, 365 days a year. This information is supplied to the National Meteorological and Hydrological Services of the organisation's Member and Cooperating States in Europe, as well as other users worldwide.

EUMETSAT is operating the geostationary satellites Meteosat-8 and -9 over Europe and Africa and Meteosat-6 and -7 over the Indian Ocean. Metop-A, the first European polar-orbiting meteorological satellite, was launched in October 2006 and has been delivering operational data since May 2007. The launch of Jason-2 in June 2008 added ocean surface topography to the missions EUMETSAT conducts. The data, products and services from EUMETSAT's satellites make a significant contribution to weather forecasting and to the monitoring of the global climate.

For further information, please visit www.eumetsat.int or contact us at EUMETSAT,
Am Kavalleriesand 31, 64295 Darmstadt, Germany
User Services: Tel +49-6151-807366/377;
Fax +49-6151-807379; E-mail ops@eumetsat.int

* pending ratification (Mar 2009)

EUMETSAT AND GMES

The European Global Monitoring for Environment and Security (GMES) initiative encompasses areas in which EUMETSAT has a clear mandate under its Convention: weather, climate change and the environment. As the organisation responsible for developing and operating Europe's operational meteorological satellites, EUMETSAT possesses the infrastructure, services and expertise required to contribute to the proposed GMES network and operate its ground systems, Sentinel satellites and future missions.
EUMETSAT maintains close links with the GMES Bureau responsible for coordinating GMES activities and the European Space Agency (ESA) to ensure that the initiative fully exploits both the investments made in meteorological satellites and EUMETSAT's capabilities and experience for the benefit of GMES users.

The exchange of letters between the European Commission and EUMETSAT and the agreement with ESA on EUMETSAT's operational role regarding Sentinel-3 (marine part), -4 and -5 are concrete steps towards the achievement of these objectives.
In the European Commission communication "GMES: We care for a Safer Planet", the role of EUMETSAT is established as the operator of the Sentinel-3 (marine part) and of Sentinel-4 and Sentinel-5 (on board EUMETSAT missions). This role has been also recognised and endorsed at the ESA Ministerial Council concluded on 26 November 2008.

EUMETSAT's CONTRIBUTION TO GMES

EUMETSAT already contributes to the objectives of the GMES programme through its mandatory and optional programmes and the delivery of data and products highly relevant to environmental monitoring.
EUMETSAT has participated in GMES activities since the launch of the initiative in Baveno in 1998. EUMETSAT's unique experience and proven capability as a provider of near-real-time data with guaranteed timelines and reliability to users 24 hours a day, seven days a week is crucial to its operational remit.

The planned EUMETSAT contribution to GMES is to provide:

- Data from all current and future EUMETSAT satellites: Meteosat , Metop , Jason, Meteosat Third Generation (MTG) , the Post-EUMETSAT Polar System (Post-EPS).
- Data and products from US National Oceanic and Atmospheric Administration (NOAA) satellites, as part of the existing programme and relations with NOAA.
- Data and products from other international satellite operators with whom EUMETSAT has signed cooperation agreements.
- Products and tools developed within the EUMETSAT ground segment, including its Satellite Application Facilities (SAFs), for a variety of operational environment monitoring activities, for example Ocean and Sea Ice SAF products for the Marine Core Service and Ozone Monitoring SAF products for the Atmospheric Services.
- Access to EUMETCast , EUMETSAT's satellite-based data distribution system.
- Support to the European Commission for the development of operational services.
- Support to GMES' international dimension by building capacity in developing countries through the Preparation for Use of MSG in Africa (PUMA).
- Support to the African Monitoring of Environment for the Sustainable Development (AMESD) initiative and to the Lisbon Process for GMES and Africa, which should plan for the extension of GMES to Africa and the African, Caribbean and Pacific (ACP) Group of States, as requested by the African and ACP Secretariat through the Maputo Declaration.

EUMETSAT AND THE CZECH REPUBLIC - The Czech Republic: Preparing to become a EUMETSAT Member State in 2010

The Czech Republic became a EUMETSAT Cooperating State in 2005 Since then, the Czech Hydrometeorological Institute (CHMI) has unlimited access to all EUMETSAT data, products and services for operational and research activities in meteorology and hydrology. These data will improve the Czech Republic's surveillance and operational forecasting and aid planning and decision-making at the national level. Once a Member State, the Czech Republic can participate fully in EUMETSAT's decision-making process, its industry can bid for EUMETSAT contracts, and Czech nationals can become members of staff.

The government of the Czech Republic has formally decided to join EUMETSAT as of 1 January 2010. Discussions on the accession process are currently ongoing.

IBM Czech Republic

IBM is the biggest IT product and services provider in the world. The company has over 80 years experience in innovations that change our world. The main line of business of IBM Czech Republic is the sale of a wide spectrum of IT technologies, from Data Storage servers and systems to software and IT services, including consultancy. Providing the complex services of a system integrator and promoting the benefits of e-commerce in everyday life belong among the main goals of IBM. Integral parts of IBM strategy are Corporate Social Responsibility programs, designed to address specific social and environmental needs. In terms of revenues, breadth of product offering and number of employees, IBM's position in the Czech environment equals its position worldwide.

IBM Global Technology Services

Services and solutions play a key role in IBM's extensive offering worldwide. IBM Global Technology Services Czech Republic focuses on providing IT services and solutions in line with the latest market trends and specific customer demands.
Thanks to acquisition of Internet Security Systems IBM can offer complete security solution, key areas also include Green Data Center solution, design, development and deployment of solutions (from data networks to applications) providing specific IT functionality to support business operations, IT infrastructure monitoring and administration services, service-level agreement (SLA) tracking and management, maintenance and technical support. Enterprise-Application Integration (EAI), Business Process Management, Automated Document Processing Systems, Disaster Recovery, Data Storage services and solutions, Business Intelligence, Data Management Services, end-to-end or partial IT takeover and management (Strategic Outsourcing) etc. are examples of other areas.
In 2001 IBM Czech Republic established a center of strategic services in Brno. Main scope of work of this center comprises remote IT system administration, end-to-end user support (helpdesk), support of network devices or computer application, or quality control management. The full scope of services is being provided by IDC Brno to clients from all over the world in many international languages.

IBM Global Financing

IBM Global Financing enables customers and business partners to acquire IT products and services using simple financial products.

IBM Global Business Services

IBM provides consulting services via its organizational unit Global Business Services which core was created by acquisition of PwC Consulting in 2002. The unit has globally over 60,000 experienced consultants and operates in over 50 countries of the world. It has grown into a unique organization with product offering that covers all consulting services - business strategy, consulting, process strategy implementations as well as IT. Portfolio of services includes Application Maintenance and Application Outsourcing, System Integration, Customer Relationship Management, Supply Chain Management, Financial Management Solutions, HR Solutions and Strategy and Change.

IBM Hardware

IBM is a traditional vendor of servers, graphic workstations, storage systems and point-of-sale systems. Implementing the latest technologies, IBM's servers achieve outstanding price/performance ratio, provide advanced virtualization features, enable "on demand" capacity expansion, thus dramatically increasing system resources utilization, performance, reliability and return on investment. IBM offers a broad range of servers from System z mainframes through Unix-based servers System p, x86 (Intel-based) servers System x, midrange servers System i to blade servers IBM BladeCenter and clusters. IBM System Storage product line includes disk arrays, tape libraries, SAN and NAS systems from entry-level models to top performing high-end systems.

IBM Software

IBM's software division holds a prominent market position with its families of:

- communication and team collaboration applications (Lotus)
- data, IT and asset management, security, IT monitoring and backup solutions (Tivoli)
- internet application development, operation and integration solutions, SOA (WebSphere)
- data storage and content management products (Information Management - DB2 & Informix)
- software development tools and project management (Rational)

In PLM area (Product Lifecycle Management), IBM Catia solutions have dominant position on the market in the long term. IBM software is of multiplatform character and fully supports Linux, open standards and open formats.

IBM Česká republika, spol. s r.o.
The Park, V Parku 2294/4
148 00 Praha 4 – Chodov
tel.: +420 272 131 111
fax: +420 272 131 401
czdirect@cz.ibm.com
ibm.com/cz

EBONE

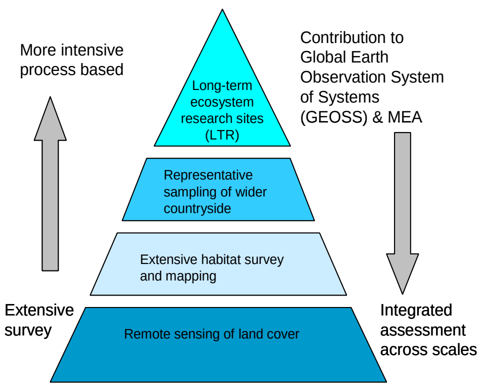
European Biodiversity Observation Network

Design of a plan for an integrated biodiversity observing system in space and time

FP7-Collaborative Project, Theme 6, Environment, Topic 4.1.1.2. Contribution to a global biodiversity observation system Project nr 21322

The main outcome: an integrated monitoring system based on key biodiversity indicators and implementation within an institutional framework operating at the European level.

Framework for a European Biodiversity Observation and Research Hierarchy



This framework will provide continued access to indicator data for CBD reporting against the 2010 target and form the basis for the continued development of a European Biodiversity Observation system.

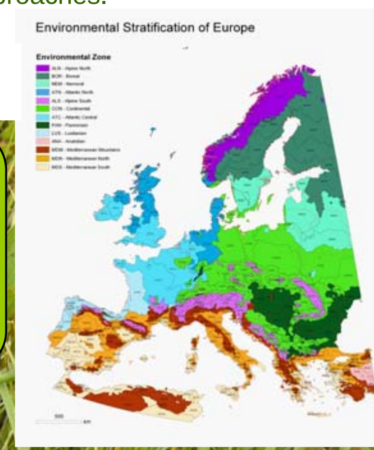
EBONE is a global pilot for international cooperation in biodiversity monitoring tackling the technical problems of harmonising approaches that differ in many ways:

Topic: species, habitats and earth observation;

Scale: from insects to migrating birds;

Biogeography: linking Boreal, Mediterranean and Desert habitats and species;

Organisation: trying to convince over 100 European agencies and an unknown number of NGOs to harmonise approaches.



Relation between working Objectives and Work Packages

WP1 Concept of monitoring, species, indicators and habitats

WP2 In situ and EO observation method of Biodiversity – overview of existing larger monitoring and surveillance systems in Europe

WP 3 Stratification and statistical analysis – creation of a statistically robust framework for monitoring

Elaboration of a monitoring concept including common indicators for biodiversity

Stratification of Europe and other regions involved for monitoring purposes

WP5 Inter-calibration of EO data with in situ observation

Development and testing of standard field-site observations and database management

Relating Earth Observation data with field data

WP4 Protocols for harmonisation of in situ data
WP6 Field validation of the methodological framework
WP 7 Data base management and IT tools

Facilitation of an institutional framework for European and world-wide biodiversity monitoring

WP8 Design of a monitoring system and its cost-effectiveness

WP9 Contribution to a world wide monitoring system: a pilot monitoring for global Mediterranean

WP10 Stakeholder involvement, communication and dissemination

Proceedings of the European conference Towards eEnvironment. Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe

Prague, March 25 – 27, 2009

Chief editor:
prof. RNDr. Jiří Hřebíček, CSc.

Published by Masaryk University, Brno, 2009
1st edition –500 issues
Number of pages: 650
Printed in DISKUS, Sokolovská 154, 180 00 Praha
Serial number Př-2/09-02/58
ISBN 978-80-210-4824-9